IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 22, 2019, accepted November 6, 2019, date of publication November 11, 2019, date of current version November 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2952564

Learning Sequence-Based Fingerprint for Magnetic Indoor Positioning System

YUANYI CHEN[1,2], MINGXUAN ZHOU[1], AND ZENGWEI ZHENG[1]
[1]Department of Computer Science and Computing, Zhejiang University City College, Hangzhou 310015, China
[2]Department of Electrical Engineering and Computer Science, University of California at Irvine, Irvine, CA 92697, USA

Corresponding author: Zengwei Zheng (zhengzw@zucc.edu.cn)

This work was supported in part by the Young Scientists Fund of the National Natural Science Foundation of China under Grant 61802343, in part by the Natural Science Foundation of Zhejiang Province under Grant LGF19F020019, and in part by the State Key Laboratory Breeding Base—Shanghai Key Laboratory of Scalable Computing and System, Shanghai Jiaotong University.

ABSTRACT Indoor magnetic-based positioning has attracted tremendous interests in recent years due to its pervasiveness and independence from extra infrastructure. Existing methods for indoor magnetic-based positioning are either point-based fingerprint matching or sequence-based fingerprint matching using the raw magnetic field strength. However, the magnetometers in smartphones are vulnerable to a few factors such as user's postures and walking speed, which causes the magnetic field strength corresponding to a location often shift in time or exhibit local distortions, thus greatly limits the positioning performance of existing methods rely on raw magnetic field strength. To this end, we observe the differences among magnetic field strength sequences are mainly attributed to small local segments, and design a new sequence-based fingerprint based on the differences among small local segments of raw MFS sequence to represent raw MFS sequence for indoor positioning. To demonstrate the utility of our proposed sequence-based fingerprint, we have performed a comprehensive experimental evaluation on two datasets, the results show that the proposed approach can significantly improve positioning performance compare with baseline methods.

INDEX TERMS Indoor positioning, magnetic field, sequence-based fingerprint, smartphone.

I. INTRODUCTION

Recent years have witnessed an increasing attention on indoor positioning in view of its importance to indoor location-based services, such as indoor advertising [1], patient activity monitoring [2] and indoor location recommendation [3]–[5]. Although a few indoor positioning methods have been proposed (e.g., UWB [6], RFID [7], Bluetooth [8], WiFi [9] and infrared-based techniques [10]), they have some inherent limitations: Ultrasound is vulnerable to indoor reflection and scattering, RFID and Bluetooth-based positioning require extra infrastructure and have small coverage range, Infrared-based positioning cannot cross walls or other obstacles, WiFi-based positioning is unstable due to heterogeneous devices and dynamic environment status.

Magnetic-based positioning becomes more and more popular in recent years since it is omnipresent and incurs almost no additional energy consumption. Researchers have found that the magnetic field strength (MFS) in indoor environments is sufficiently stable, i.e., the variation of MFS at a location over time is negligible. Moreover, the values of MFS at different positions are different and consequently magnetic-based positioning is possible by pattern matching of the MFS records. In addition, the MFS is robust to indoor multipath phenomena and without LOS operating conditions. A few magnetic-based positioning methods [11]–[21] have been proposed in the literatures due to these special properties (for a review sees Section 2). However, existing magnetic-based positioning methods suffer from two challenges. First, existing point-based fingerprint matching using the 3-D MFS vector makes no sense due to the following two factors: 1) the MFS at a given indoor location is a 3-D vector in space that varying similarly with near location; 2) different orientation or postures of mobile phone lead to different MFS readings at the same location. Second, existing sequence-based fingerprint matching is far from satisfactory since the MFS values measured by mobile devices are vulnerable to external magnetic perturbations (e.g., the MFS may be distorted by metal or noise sources deriving from the surrounding environment), which often result in a magnetic distortion known as soft and hard iron effects.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhibo Wang.

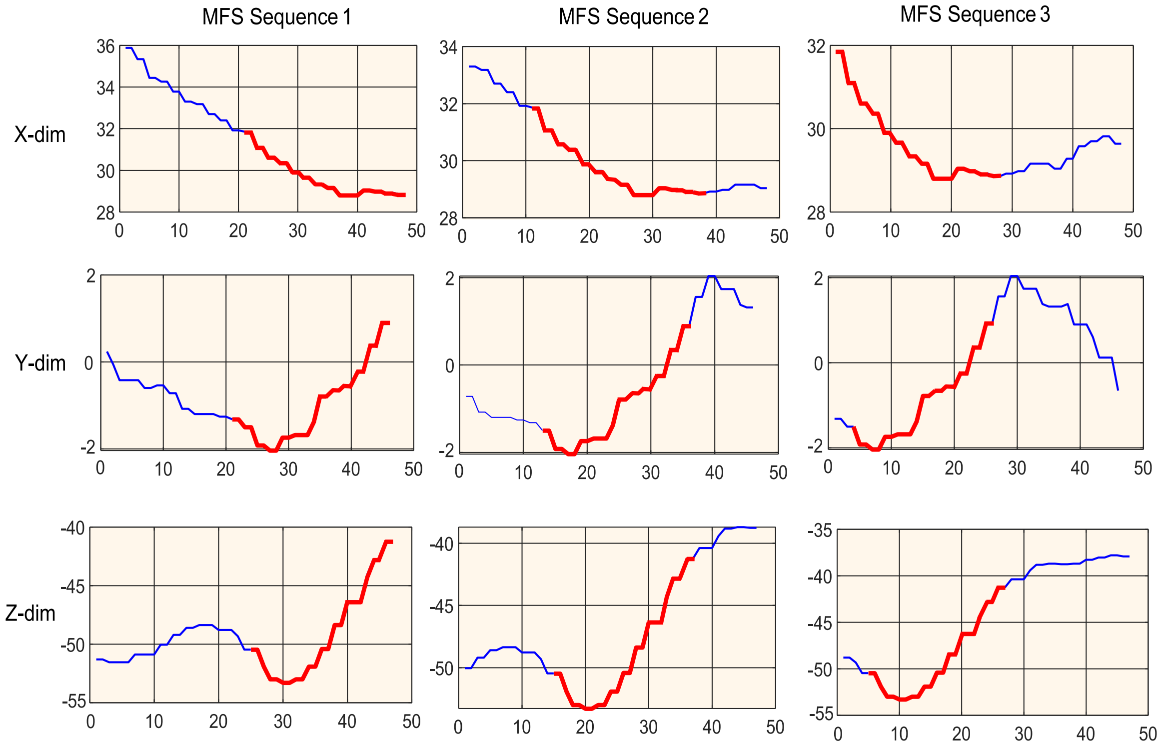

FIGURE 1. Three MFS sequences of the same space trajectory and the discriminative segment is highlighted with red color for each sequence.

Even though a number of studies have been proposed to address the challenges of magnetic-based positioning recently [22]–[30], these approaches also have a number of limitations. For example, the studies in [22]–[24] tackled the two challenges by incorporating particle filters along with inertial sensors. Unfortunately, they cannot obtain reliable positioning results because dead reckoning relies on inertial sensors will suffer from serious cumulative error, and continually collecting data from multi-sensor is energy-consuming. Another literatures [25]–[28] addressed the challenges with hybrid positioning by jointly utilizing magnetic field and other information (e.g., WiFi, channel state information and visual images). However, WiFi radio signal strength (RSS) easily fluctuates due to the multipath effect such as the diffraction, scattering and reflection in indoor environments. Although channel state information is fine-grained value from the physical layer compares with WiFi RSS, but it requires customized equipment. Hybrid positioning using visual images and magnetic field is time consuming and energy-consuming. Still other researchers solved the challenges of localization systems with magnetic field by recognizing the magnetic sequence pattern, but most attempts such as dynamic time warping (DTW) [29] and bag of words [30] cannot effectively capture the pattern in raw MFS sequence. The reason is even along the same path, the magnetic intensity on the path fluctuates with a slight change in location due to a few factors, such as user's walking speed and the device's height.

Although magnetic sequence pattern might be shifted in time and exhibit local distortions/noise, we observe the pattern might be present on short local segments of raw MFS sequence rather than on its global structure. For instance, three MFS sequences for the same space trajectory are shown in Figure 1, we observed that although the three sequences are distorted in temporal and spatial, the discriminative patterns in short segments are obvious and stable for each MFS sequence. On this basis, we propose an indoor magnetic-based positioning system by learning sequence-based fingerprint from local raw MFS sequences in this paper. Specifically, we firstly learn sequence-based fingerprint with high discriminative power using training data. Then, we transform the raw MFS sequences of test dataset into the learnt fingerprint representation and estimate the unknown location of a test sample by training a classification model. To summarize, the main contributions of this paper are as following:

- We focus on indoor magnetic-based positioning and propose a novel sequence-based fingerprint based on the differences among small local segments of raw MFS sequence, which can efficiently handle the local distortions and shift of raw MFS sequence.
- Different from taking raw MFS sequence as features directly, we firstly learn sequence-based fingerprint with high discriminative power using training data, to estimate the unknown location of a test sample by training a classification model.
- We conducted extensive experiments on two real-world datasets, and the experimental results on the task of indoor localization show that our proposed approach significantly outperforms existing

state-of-the-art approaches in terms of both mean positioning error and cumulative error distribution.

The remainder of the paper is organized as follows: Section II surveys related work on indoor positioning with magnetic field. Section III describes the proposed indoor magnetic-based positioning model in detail. Section IV reports and discusses the experimental results. Finally, we present our conclusion and future work in Section V.

II. RELATED WORK

A few magnetic-based positioning techniques [11]–[21] have been proposed due to the stability and uniqueness of the magnetic field. The work in [11] proposed an energy-efficient indoor positioning method based on MFS sequence with an improved Dynamic Time Warping algorithm. Gozick *et al.* [12] proposed a magnetic map construction method by building the space magnetic distribution using an magnetism calculation technique. However, the proposed positioning system is impractical since it utilizes the magnetic anomalies around pillars that may be rare in some indoor environments. MaLoc [13] utilized magnetic magnitude difference as fingerprint to search real-time sampling data in magnetic maps. Their results reveal magnetometer's sensitivity is different for different devices. Meng *et al.* [14] constructed magnetic map for online positioning using local weight regression. The study [15] proposed a positioning method based on low-frequency magnetic that exploited the principle of inductive coupling between tuned wire loop antennas. Chung *et al.* [16] proposed a positioning method with an array of magnetometers by measuring the magnetic readings of all directions at a position. The improved work [17] utilized all three coordinates of magnetic field to build localization fingerprint instead of the commonly used magnitude, which can avoid the measurements for all directions by coordinate transformation, but it is error-prone since the orientation estimation usually contains errors. In [18], many magnetic field fingerprint features (such as kurtosis, mean and slope) are tested and can achieve room-level accuracy. The work [19] studied the magnetic field intensity and direction distribution features for constructing magnetic maps of indoor space. To improve the positioning accuracy, this study [20] using two kinds of automatic calibration methods, i.e., opportunistic magnetic trajectory matching and indoor landmark identification. The study [21] performed indoor magnetic-based localization by using deep learning to recognize magnetic sequence patterns. However, these attempts are proven only in an one-dimensional space, because magnetic intensity fluctuates severely with even a slight change of locations.

Unfortunately, the studies [31]–[33] demonstrated that the 3-D magnetic fingerprints may not be unique in a large indoor space. In addition, different orientations of mobile devices lead to different readings since the metal construction of magnetometer could influence the magnetic field. In addition, sensor noise is unavoidable primarily due to hasty movements of a smartphone user and the inherent bias of different smartphone sensors. These factors make positioning systems with raw magnetic field strength may achieve poor positioning performance.

Recently, researchers proposed various new techniques to address the challenge of indoor positioning based on raw magnetic field strength. For example, the studies in [22]–[24] utilized particle filter to tackle these challenges and improve positioning accuracy. The basic principle is to model user's state such as position and heading direction as particles and represent the posterior distribution of a user's state. More exactly, the work [23] designed a technique for map matching where the pedestrian movement is matched to a map of the magnetic landscape based on particle filter. However, existing motion estimation methods are error-prone when applying to smartphone tracking, motion estimation usually incurs much more noise than robots thus greatly limit the positioning accuracy. To minimize errors in motion estimation and improve the robustness of the basic particle filter, the study [24] proposed reliability-augmented particle filter for magnetic-based positioning based on dynamic step length estimation and heuristic particle resampling. However, all particles may lost tracking of the user and is unrecoverable in highly noisy environments. To solve this, the work [22] proposed a hybrid positioning by fusing geomagnetic and visual sensing, used a context-aware particle filtering framework to track the user with the goal of maximizing the positioning accuracy.

Other studies [25]–[28] address the challenge of indoor positioning with raw magnetic field by integrating information from other sources. The works [25], [26] incorporated WiFi signals to achieve much improved positioning accuracy for indoor environments. But the WiFi RSS strength easily varies due to the multipath effect (e.g., diffraction, scattering and reflection in indoor environments. To compensate for the multipath defect of WiFi RSS, the study [27] utilized channel state information to reduce the localization error of magnetic fingerprint positioning. A few studies [28] fused geomagnetic and visual sensing to improve the positioning accuracy. However, indoor positioning by fusing magnetic field and other information obviously incurs high training cost and may need extra infrastructure.

Our proposed approach differs from the above-mentioned work in the following two aspects: 1) the proposed approach performs indoor positioning based on magnetic field solely, which is scalable and pervasively available. 2) the proposed approach learns a set of sequence-based fingerprints with high discriminative power from labeled magnetic sequences, which can boost indoor localization accuracy. In addition, the proposed approach is energy-efficient since the learnt sequence-based fingerprint is usually much smaller than the raw magnetic sequence.

III. THE PROPOSED INDOOR POSITIONING MODEL

In this section, we first present the problem statement of indoor positioning with MFS sequence data. Then detail the

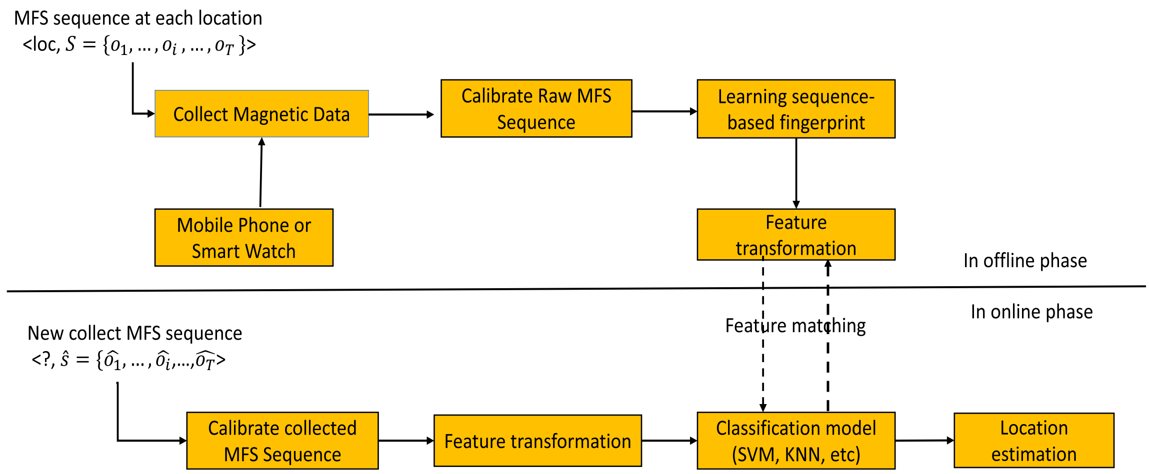

FIGURE 2. The framework of the proposed indoor magnetic-based positioning method.

TABLE 1. Notations used in the paper.

SYMBOL	DESCRIPTION
N,M,K	The number of D_{tr},D_{te} and C
D_{tr}	Train dataset that contains a few MFS sequences labeled with space trajectories
D_{te}	Test dataset that contains a few MFS sequences
C	The set of space trajectories in indoor environment
$o(u,t,R)$	MFS sample collected by user u at time t
S	A MFS sequence that contains a set of MFS samples
S_p^l	A MFS subsequence of S that contains l continuous MFS samples and starts at position p
$subDist(S_p^l,S)$	The minimum distance between MFS subsequence S_p^l and MFS sequence S
$\widehat{A}_i$	A MFS sequence need to find the corresponding space tarjectory

proposed approach, a fingerprint-based positioning method by learning sequence-based fingerprint based on the differences among small local segments of raw MFS sequence.

A. PROBLEM STATEMENT

For ease of the following presentation, we define the key data structures and notations used in the proposed method. Table 1 lists the relevant notations used in this paper.

Definition 1 (MFS Sample): A MFS sample is a triple $o(u, t, R)$ that means the collected magnetic field strength by user u at time t. R is a 3-D vector and denoted by (r^x, r^y, r^z). Actually, the magnetometer measures the value of the magnetic field strength at a position in space relative to the orientation of the smart phone. This magnetic field strength is a vector of three coordinates, each one representing the value of the magnetic field along one direction of the phone's frame (x, y, z).

Definition 2 (MFS Sequence): We define a MFS sequence as a set of MFS samples and denoted by $S = \{o_1, \ldots, o_i, \ldots, o_T\}$, o_i represents the collected MFS sample at time t_i, $1 \leqslant i \leqslant T$. Typically, The MFS sequence is utilized to predict a short indoor trajectory and has about 50 discrete MFS samples [34].

Definition 3 (MFS subsequence): A MFS subsequence S_p^l is a set of continuous MFS samples from MFS sequence S, which starts at position p and the number of MFS samples is l.

Definition 4 (SubSequence Distance): SubSequence Distance $subDist(S_p^l, S)$ represents the minimum distance between S_p^l and all the subsequences of S with length l, denote by:

$$subDist(S_p^l, S) = min\{dist(S_p^l, S_1^l), \ldots, dist(S_p^l, S_{L-l+1}^l)\} \tag{1}$$

where $dist(S_{L-l+1}^l, S_p^l)$ is the Euclidean distance between S_{L-l+1}^l and S_p^l, l and L are the length of subsequence S_p^l and sequence S respectly.

Definition 5 (Train Dataset): A train dataset D_{tr} is a collection of N MFS sequences $\{S(1), S(2), \ldots, S(N)\}$ labeled by K space trajectories $C = \{c_1, \ldots, c_i, \ldots, c_K\}$. Thus, D_{tr} is a set of pairs of MFS sequence and the corresponding space trajectory: $D_{tr} = \{(S(i), c_k) : 1 \leq i \leq N, 1 \leq k \leq K\}$.

Based on the above definitions, we formulate the problem statement of indoor magnetic-based positioning as: Given: 1) Train dataset $D_{tr} = \{(S(i), c_k) : 1 \leq i \leq N, 1 \leq k \leq K\}$ generated by manual annotation; 2) A few MFS sequence $D_{te} = \{A_1, \ldots, A_i, \ldots, A_M\}$ collected by users who need for positioning; Objective: find the space trajectory in C that corresponding to $A_i (1 \leq i \leq M)$ by MFS sequence matching.

Basically, one can find the space trajectory corresponding to A_i by matching MFS sequences with a similarity measure, such as Euclidean distance and dynamic time warping. However, the raw MFS sequence of the same space trajectory collected by users may be distortions in time and exhibit local noise, similarity measure relies on raw MFS sequences cannot find the correct space trajectory. To tackle this challenge, we firstly extract sequence-based fingerprint with high classification accuracy from the raw MFS sequences, then matching MFS sequences by training a classification model based on the learnt sequence-based fingerprint.

B. THE PROPOSED APPROACH

As shown in Figure 2, the proposed model performs indoor positioning by two phases: 1) offline learning sequence-based fingerprint via labelled MFS sequences and building localization map of indoor space; 2) Online estimating the unknown position with a classification model. In online localization phase, the test MFS sequences are first transformed with the learnt sequence-based fingerprints, then are fed into a classification model for location estimation. Note that the raw MFS Sequence needs to be calibrated since many factors can influence the MFS strength during data collecting.

1) CALIBRATING RAW MFS SEQUENCE

A few factors can influence the MFS records during collecting, such as mobile phone pose, metallic materials and noise in magnetometer measurement, thus the MFS sequences are pretreated to remove the added noise for improving the positioning accuracy. Existing studies [23], [24] demonstrated the MFS records should change smoothly in a small continuous area. Therefore, the MFS records fluctuate significantly within a small area can be utilized to identify noise data. On the basis, we eliminate noise data by two stages: 1) Identify all noise data in raw MFS sequence by a sliding window-based technique; 2) Replace noisy data with predicted values of the autoregressive model.

Step 1: Identify Noise Data in Raw MFS Sequence. We utilize the fluctuation degree of MFS records in a small sliding window to identify noise data. Formally, given a MFS sequence $S = \{o_1, \ldots, o_i, \ldots\}$, we define $Var(t_i, r^d, \tau)$ to represent the d-dimension fluctuation degree of MFS sequence in sliding window $(t_i - \tau/2, t_i + \tau/2)$, as shown in Equation 2.

$$Var(t_i, r^d, \tau) = \frac{1}{\tau - 1} \sum_{j=t_i-\tau/2}^{t_i+\tau/2} (r_j^d - \overline{r_j^d})^2 \tag{2}$$

where $\overline{r_j^d}$ is the average MFS values from d-direction in time window $(t_i - \tau/2, t_i + \tau/2)$, r_j^d is the MFS value from d-direction at time t_j.

If the MFS fluctuation degree in time window $(t_i - \tau/2, t_i + \tau/2)$ is significantly higher than average, we can infer it is a noise data at time t_i. Formally, we use variation coefficient α to quantify the fluctuate degree, as shown in Equation 3.

$$\alpha = \frac{|S| \times Var(t_i, r^d, \tau)}{\sum_{j=1}^{|S|} Var(t_j, r^d, \tau)} \tag{3}$$

Algorithm 1 Learning Sequence-Based Fingerprint of Training Dataset

Require: 1) $D_{tr} = \{S(1), S(2), \ldots, S(N)\}$; 2) the minimum and maximum length of sequence-based fingerprint: *minL* and *maxL*; 3) the number of learnt sequence-based fingerprints: k.

Ensure: The learnt top-k sequence-based fingerprints: $U_k = \{u_1, u_2, \ldots, u_k\}$

$U_k = \emptyset$;
for each MFS sequence $S(i) \in D_{tr}$ **do**
  $\varphi = \emptyset$;
  for l = *minL* to *maxL* **do**
    Generate all subsequences with length l of S_i: $S(i)^l$;
    for each subsequence $sq \in S(i)^l$ **do**
      Dist(sq) = subDist(sq,S(i));
      score = evaluateCandidate(S(i),sq);
      Add $< S_i, sq, score >$ to φ;
    end for
  end for
end for
Sort φ by score and remove self-similar subsequences;
return Select top-k subsequences in φ and add to U_k;

For example, set time window size $\tau = 5$ and variation coefficient as $\alpha = 3$, the three direction's fluctuation degree from MFS sequence in Figure 3a is calculated as shown in Figure 3b. As shown in Figure 3c, we infer the segments in the red dotted box may contain noise data.

Step 2: Replace Noise Data with Predicted Values of Autoregressive Model. We infer a sliding window may contain noise data using a threshold-based technique in Step 1, then detect the exact location of the noise data based on an autoregressive model. The principle is the difference between the forecasted value and the actual value can be calculated as the noise score, which represents a deviation between the expected normal behaviour and actual behaviour. Formally, given a MFS sequence that may contain noise data in sliding window $\{x_1, x_2, \ldots, x_\tau\}$, the MFS value in the sliding window is forecasted by Equation 4.

$$\overline{x_t} = -\sum_{i=1}^{\tau} a_i x(t-i) \tag{4}$$

The autoregressive model aims to find a_i, i.e., the autoregressive coefficients, to minimize the squared error between x_t and $\overline{x_t}$. The autoregressive coefficients can be estimated by maximum likelihood estimation. Once we have a forecasted value from autoregressive model for an incoming actual value, the noise data are replaced by the forecasted value.

2) LEARNING SEQUENCE-BASED FINGERPRINT FROM LABELED MFS SEQUENCE

In this study, the indoor positioning problem is formulated as predicting the space trajectory of user-generated MFS

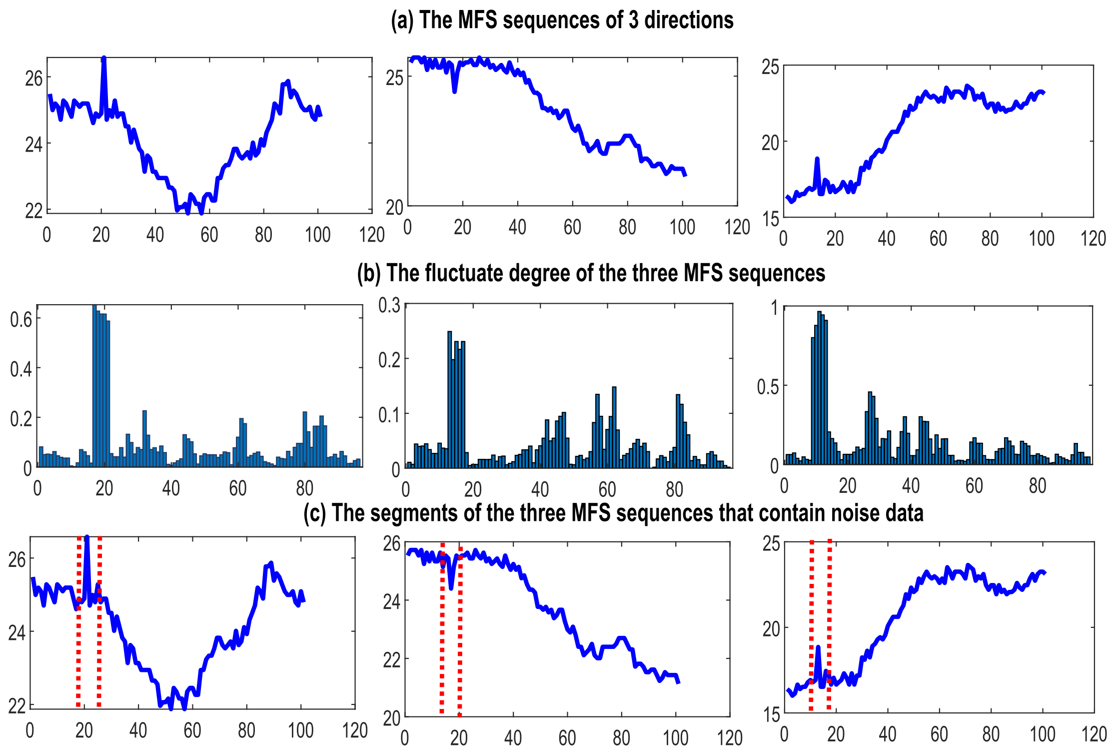

FIGURE 3. An example of identifying noise data in raw MFS sequence.

sequences based on a few labeled samples. However, the raw MFS sequence might be shifted in time and the difference might be present on short local segments rather than on global structure, as shown in Figure 1. Furthermore, we observe most MFS sequences have the same global structure but differ only in the highlighted segments. On the basis, we perform MFS sequence matching by learning sequence-based fingerprint of raw MFS sequence with high discriminative power. Specifically, the sequence-based fingerprint of a raw MFS sequence is a set of its local subsequences, which is evaluated from numerous candidate subsequences based on shapelet discovery [35].

Formally, the sequence-based fingerprints are a set of $u \in U$ patterns and each having length $h \in H$, and denoted as $U \in R^{U \times H}$. We consider all possible subsequences of a kind of MFS training series as potential candidate sequence-based fingerprints. The minimum distances between sequence-based fingerprint $U_{u,h}$ and all the subsequences of a MFS sequence S were used as a feature for ranking the information gain accuracy of that candidate on the target MFS sequence, the minimum distance of a set of sequence-based fingerprints to MFS sequences can be perceived as a kind of data transformation, namely the shapelet-transformation representation (more details about the shapelet-transformation representation can be found in [36] and [37]), denoted as $subDist(S, U_{u,h})$ as shown in *Definition 4*. The challenge of this representation is to find the sequence-based fingerprints U, for which the resulting representation helps achieve the highest classification accuracy.

We learn top-k sequence-based fingerprints of training dataset D_{tr} by the following four steps, as shown in Algorithm 1 (see Figure 4):

(a) **Candidate Fingerprints Extraction**: Iterating over all the subsequences of the train dataset D_{tr} and considers each subsequence as a candidate sequence-based fingerprint, as shown in Line 2∼6 of Algorithm 1;

(b) **Candidate Fingerprints Scoring**: For each candidate fingerprint, the raw training set is transformed by calculating the minimal Euclidean distance between a training sample and the candidate fingerprint. Then, utilizing the minimal Euclidean distance to the samples as a classification feature and evaluating the candidate fingerprint using F-stats measure [36], as shown in Line 7∼9 of Algorithm 1;

(c) **Sequence-based Fingerprints Discovery**: Sorting the candidate fingerprints based on F-stats score and selecting the top-k candidates with the highest classification accuracy as the utilimate sequence-based fingerprints, as shown in Line 13∼14 of Algorithm 1.

3) INDOOR MAGNETIC-BASED POSITIONING BASED ON THE LEARNT SEQUENCE-BASED FINGERPRINT

As shown in Figure 5, the collected MFS sequence $S = \{o_1, \ldots, o_i, \ldots, o_T\}$ are firstly calibrated when user sends localization request. Then, we transform both the raw train datasets and test MFS sequence into a new feature representation by calculating the minimum distance to the learnt

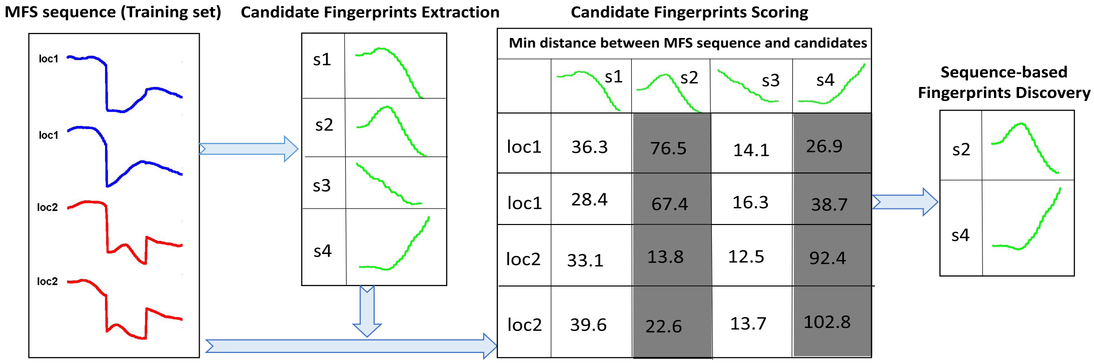

FIGURE 4. An illustration of learning top-k sequence-based fingerprints.

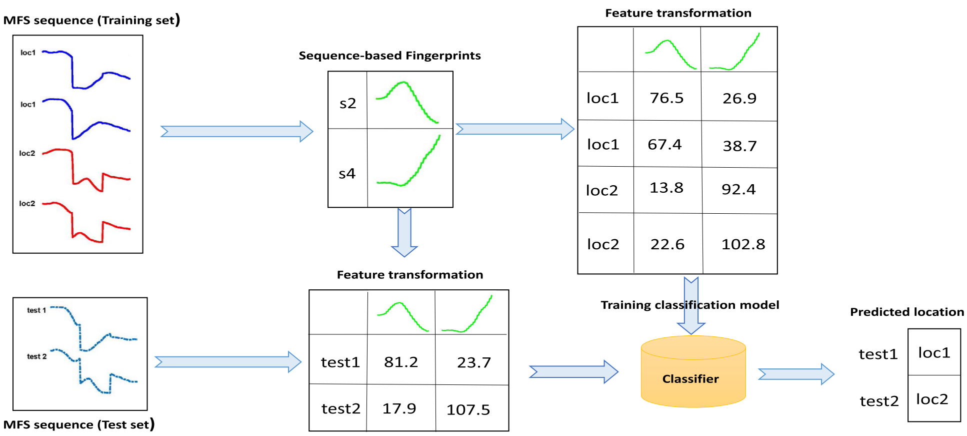

FIGURE 5. An example of online localization process.

top-k sequence-based fingerprints, and perform dimension reduction using Principal Component Analysis. Finally, our method estimates the positioning of user's current location by training a classification model (e.g., KNN, Random Forest and SVM) using the transformed feature representation.

IV. EXPERIMENT EVALUATION

In this section, we report on the results of a series of experiments conducted to evaluate the performance of the proposed indoor magnetic-based positioning model. We first describe the settings of experiments including data sets, comparative methods and evaluation metric. Then, we report and discuss the experimental results.

A. EXPERIMENTAL SETTINGS

1) DATA SETS

Two real world datasets are used for experimental evaluation in this study : 1) a large-scale public benchmark dataset, namely *UJIIndoorLoc-Mag* [34], is used to compare the proposed method with some state-of-the-art methods; 2) one collected dataset in our experiment, is used to verify the effectiveness of the proposed method in different test situations (such as different collecting heights, devices and collecting times.

- **UJIIndoorLoc-Mag:** the dataset is published by university Jaume I in the 2015 International Conference on Indoor Positioning and Indoor Navigation, which is collected from a 260 m^2 laboratory with 8 corridors(please refer to [34] for the data collection environment and test trajectories). For training part, the data is collected by repeating 5 times from 54 different alternative paths, thus 270 different continuous MFS sequences are used to generate training dataset. For testing part, the data is collected on 11 trajectories that contain 9 complex routes. Each MFS sequence of training and test dataset is divided in several subsamples with 5 seconds for each one. More details of the dataset are reported in Table 2 and Table 3. For point-based fingerprint matching, the discrete samples are extracted from the continuous training MFS sequences, while the samples from the 11 test paths are used as test dataset. Each sample for point-based fingerprint matching consists of 5 features: the Latitude and longitude where the sample was taken $[lat, lon]$ and the MFS values from three directions in this location (r^x, r^y, r^z). In total, the training dataset includes 8943 discrete samples and the test dataset includes 4380 discrete samples.
- **Our dataset:** Our data collection environment includes three corridors and two classrooms. Three volunteers

TABLE 2. The number of samples for each test trajectory.

	1	2	3	4	5	6	7	8	9	10	11
Continuous MFS sequences	35	21	44	41	23	9	8	16	11	10	13
Discrete samples	540	356	876	859	362	224	211	246	196	223	287

TABLE 3. Statistics of indoor positioning datasets.

	#space trajectories	#length of samples	#samples
Training Dataset	54	50	540
Test Dataset	17	50	231

who are mainly students and different smartphones including Xiaomi 5S, Huawei Mate 10 and Vivo Y87, are used to collect dataset. For recording all the traces and the ground truth, we customize an ink stamper on each volunteerâĂŹs shoes by following prior work [13], thus can leave a stamp on the floor surface when a volunteer walking. We repeat the data collection 3 times at different dates. For each time, the training data is collected from 20 different trajectories with three different heights (0.2m, 1m, 1.8m) by each volunteer, thus 180 different continuous MFS sequences are collected each time. For testing data, the data is collected on 10 trajectories with the same three heights (0.2m, 1m, 1.8m) by each volunteer, thus 90 different MFS sequences are collected each time. Therefore, for sequence-based fingerprint matching, we collect 540 MFS sequences for generating training dataset and 270 MFS sequences for generating test dataset. The discrete samples are extracted from the continuous MFS sequences for evaluating point-based fingerprint matching. Each sample for point-based fingerprint matching consists of 5 features: the geometric coordinates where the sample was taken $[locX, locY]$ and the MFS values from three directions in this location (r^x, r^y, r^z).

2) COMPARATIVE METHODS

We compare the proposed indoor magnetic-based positioning model with the following state-of-the-art competitors, where the first method is well-known existing methods for indoor magnetic-based positioning using point-based fingerprint matching, the next three methods are start-of-art methods for indoor magnetic-based positioning using sequence-based fingerprint matching.

- **Point-based K-nearest Neighbors Algorithm (PKNN)**: the method [34] estimates the location of each test sample with k-nearest neighbors algorithm. Specifically, the location of the most similar sample in the training set is the one assigned to the test sample, and Euclidean's distance is used as similarity metrics between two samples.
- **Sequence-based Bag of Words (SBOW)**: Inspired in the bag of words representation, this method [30] firstly constructed positioning fingerprint using a simplified bag of words methodology allowing user speed invariance, then utilized classification model such as KNN and SVM to estimate the location of test sample.
- **Sequence-based K-nearest Neighbors Algorithm (SKNN)**: this method [34] proposed a MFS sequence-based fingerprint matching with KNN, in which the Euclidean's distance is used to measure the distance between two MFS sequence data.
- **Sequence-based Dynamic Time Warping (SDTW)**: this method [29] utilized a dynamic time-warping-based approach to measure the similarity between two MFS sequences generated by walking short distances, then 1-NN is used to estimate the location of test sample.

3) EVALUATION METRIC

For *UJIIndoorLoc-Mag* dataset, this distance $dist(i, j)$ between ground truth location (lat_i, lon_i) and the predicted location (lat_j, lon_j) is calculated by the haversine formula [34].

For positioning model using point-based fingerprint matching, the mean positioning error is calculated by Equation 5:

$$PErr = \frac{1}{|D_{te}|} \sum_{i \in D_{te}} dist(i, \hat{i}) \tag{5}$$

where D_{te} is the test dataset and $|D_{te}|$ is the num of discrete samples in D_{te}, $dist(i, \hat{i})$ denotes the distance between the ground truth location and the corresponding predicted location.

For positioning model using sequence-based fingerprint matching, the mean positioning error is calculated by Equation 6:

$$SErr = \frac{1}{M} \sum_{i=1}^{M} SDidst(S_j, \hat{S}_j) \tag{6}$$

where $|M|$ is the num of test MFS sequences, $SDidst(S_j, \hat{S}_j)$ denotes the positioning error between MFS sequence S_j and its predicted MFS sequence $\hat{S}_j$, which is calculated by Equation 7:

$$SDidst(S_j, \hat{S}_j) = \frac{\sum_{j=1}^{|S_i|} dist(j, \hat{j})}{|S_j|} \tag{7}$$

For our dataset, the positioning error of point-based fingerprint matching is calculated as the Euclidean distance between ground truths and estimated locations in trajectories. For positioning model using sequence-based fingerprint matching, the mean localization error is calculated as the mean of all the positioning errors along a test MFS sample.

TABLE 4. Parameter tuning for removing noisy data of MFS values.

α \ τ	4	5	6	7	8	9	10	11	12
5	3.938	3.889	3.911	3.855	3.941	3.883	3.832	3.833	3.921
6	3.896	3.853	3.888	3.931	3.855	3.909	3.856	3.879	3.847
7	3.951	3.915	3.886	3.89	3.896	3.928	3.885	**3.78**	3.851
8	3.949	3.871	3.833	3.899	3.906	3.916	3.925	3.903	3.903
9	3.870	3.956	3.938	3.876	3.878	3.922	3.919	3.875	3.859
10	3.821	3.899	3.956	3.905	3.9	3.889	3.931	3.961	3.929
11	3.887	3.852	3.915	3.903	3.847	3.903	3.927	3.903	3.878
12	3.873	3.903	3.857	3.959	3.928	3.93	3.896	3.908	3.949
13	3.895	3.905	3.903	3.927	3.938	3.906	3.825	3.898	3.901
14	3.934	3.895	3.884	3.882	3.912	3.906	3.85	3.881	3.858
15	3.962	3.907	3.85	3.936	3.918	3.908	3.881	3.827	3.894
16	3.864	3.838	3.869	3.866	3.868	3.888	3.876	3.832	3.829
17	3.897	3.908	3.879	3.871	3.833	3.906	3.978	3.822	3.935

B. EXPERIMENTAL RESULTS

We conduct three experiments on *UJIIndoorLoc-Mag*: 1) perform parameter tuning for calibrating the raw MFS sequence by a sliding window-based technique; 2) compare the positioning performance of four classification models (KNN, SVM, Random Forest and Naive Bayes) with the raw MFS sequence and the learnt sequence-based fingerprint (as introduced in Section 3.3.2); 3) compare the positioning performance of the proposed positioning model with four state-of-the-art competitors. With the collected dataset, we further evaluate the performance of the proposed method in different situations.

1) EXPERIMENTAL RESULTS OF UJIINDOORLOC-MAG DATASET

a: IMPACT OF MODEL PARAMETERS

Based on the learnt sequence-based fingerprint, 1-NN is used to estimate the location of test MFS sequence. We present the results of mean positioning error in Table 4 by varying the sliding window size τ from 4 to 12 and the variation coefficient α from 5 to 17. For each parameter setting, we repeat the process 5 times and report the average positioning error of the 5 experiments. From Table 4, we observe the positioning error is between 3.8 m to 4 m, which suggests the proposed method is not sensitive to different parameter setting. We further observe the best positioning performance is 3.78 m when setting the sliding size as 11 and the variation coefficient as 7.

b: INDOOR MAGNETIC-BASED POSITIONING WITH THE LEARNT SEQUENCE-BASED FINGERPRINT

In this part, we compare the performance of models utilizing the raw MFS sequence and the learnt sequence-based fingerprint for indoor positioning. We evaluate the positioning effectiveness by comparing the mean positioning error and the cumulative distribution function (CDF) of four positioning models (KNN, Random Forest, SVM and Naive Bayes) with the raw MFS sequence and the sequence-based fingerprint (SBFP), as shown in Table 5 and Figure 6. For each case, we report the average performance by repeating the experiments 5 times.

TABLE 5. Mean positioning error comparison with raw MFS sequence (RAW) and the learnt sequence-based fingerprint (SBFP).

	KNN	Random Forest	SVM	Naive Bayes
RAW	6.71 m	6.73 m	6.76 m	6.15 m
SBFP	3.78 m	4.07 m	4.84 m	5.98 m

From Table 5, we can observe: 1) All the four positioning models using the learnt sequence-based fingerprint always outperforms using the raw MFS sequence. For example, the mean positioning error of KNN using sequence-based fingerprint (KNN+SBFP) is reduced by more than 43% compares with using the raw MFS sequence (KNN+RAW), i.e., dropped from 6.71 meters to 3.78 meters. The results suggest that, the learnt sequence-based fingerprint can better represent the inherent pattern of raw MFS sequences thus is beneficial for improving the performance of positioning, while the raw MFS sequences are vulnerable to a few factors such as user's postures and walk speed thus achieves much worse performance; 2) For positioning models with raw MFS sequence, Naive Bayes achieves the best performance about 6.15 meters, but the performance improvements with sequence-based fingerprint (only 0.17 meters) are minimal than the other three models. For positioning models with sequence-based fingerprint, KNN achieves the best performance about 3.78 meters.

Figure 6 shows the cumulative distribution function (CDF) of positioning error by the four positioning methods with raw MFS sequence and the learnt sequence-based fingerprint, respectively. In this figure, we can see that all the positioning models with the proposed sequence-based fingerprint outperforms using raw MFS sequence. More exactly, KNN with the proposed sequence-based fingerprint (KNN+SBFP) achieved the best positioning performance and the second is Random Forest with sequence-based fingerprint (Random Forest+SBFP), while the worst performance is achieved by SVM with the raw MFS sequence (SVM+RAW). For instance, the probability of error distance under 5 meters is 58.7% by KNN+SBFP, 21.6% by KNN+RAW, 48.9% by Random Forest+SBFP and 20% by SVM+RAW. In Figure 6d, the CDF curves of Naive Bayes with raw MFS

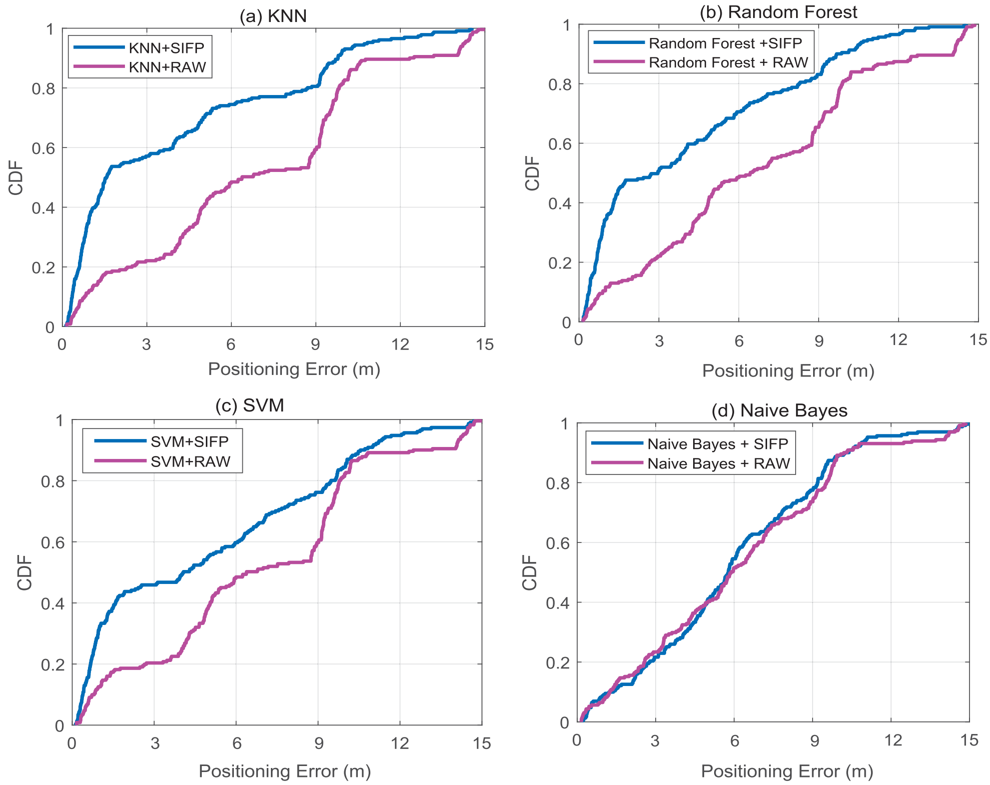

FIGURE 6. The cumulative error distribution with the raw MFS sequence and the sequence-based fingerprint.

TABLE 6. Mean positioning error compare with benchmark methods.

PKNN	SBOW	SKNN	SDTW	Our method
7.23 m	6.02 m	6.71 m	6.01 m	**3.78 m**

sequence and the learnt sequence-based fingerprint are very close, showing the performance improvement with the learnt sequence-based fingerprint is not significant for Naive Bayes model.

c: INDOOR POSITIONING PERFORMANCE COMPARISON

To investigate the advantage of learning sequence-based fingerprint for indoor magnetic-based positioning, we further compare the proposed positioning methodology (KNN+SBFP) with four baseline methods, the results are shown in Table 6 and Figure 7.

From Table 6, we observe: 1) KNN with the learnt sequence-based fingerprint achieves the best positioning performance and the mean positioning error is 3.78 meters, showing again the learnt sequence-based fingerprint can reveal the pattern of MFS values in a space trajectory; 2) the positioning methods based on MFS sequence (e.g., SBOW, SKNN and SDTW) outperform KNN using point-based fingerprint matching (PKNN). More exactly, the mean positioning error is 6.02 meters for SBOW, 6.71 meters for SKNN, 6.01 meters for SDTW and 7.23 meters for PKNN. The results demonstrate sequence-based fingerprint matching

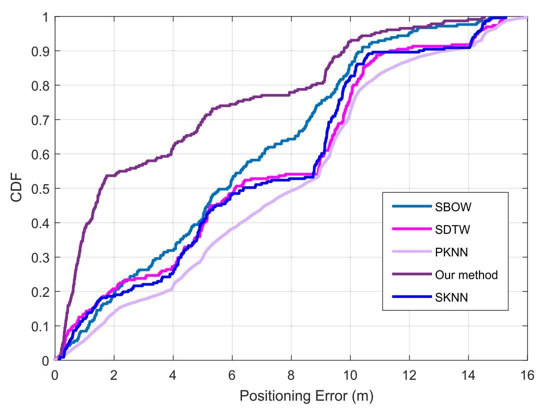

FIGURE 7. The cumulative error distribution of different positioning models.

is better than point-based fingerprint matching for indoor magnetic-based positioning, and it is sufficient to walk a short distance in indoor space to be located using the magnetic field of smartphone.

The cumulative distribution function (CDF) of positioning error for the four baseline methods and our method are shown in Figure 7. Clearly, the proposed method outperforms other baseline algorithms significantly, showing again the advantage of learning sequence-based fingerprint from the raw MFS sequences can allow more accurate indoor positioning. For instance, the probability of error distance under 4 meters is 62.1% by our method, 32.6% by SBOW, 27.7% by SDTW, 26.1% by SKNN and 21.8% by PKNN. The results

TABLE 7. Mean positioning error (m) in the 11 testing trajectories.

Method \ Trajectory	1	2	3	4	5	6	7	8	9	10	11
PKNN	8.81	7.12	7.81	7.81	6.11	7.5	7.72	9.3	3.33	7.48	6.66
SBOW	5.27	5.82	5.29	7.27	4.99	5.52	5.7	6.52	7.96	5.91	5.99
SKNN	8.64	7.35	8.08	7.32	5.28	**5.07**	6.35	6.93	**1.28**	4.92	4.12
SDTW	8.34	7.47	7.75	7.52	**4.7**	5.66	7.1	6.73	1.34	**3.87**	5.62
Our method	**2.14**	**3.39**	**2.84**	**3.29**	8.22	5.5	**5.3**	**1.75**	4.85	5.61	**3.74**

TABLE 8. The parameters of a test sample extracted from trajectory 5 and it's candidate.

	C1	C2	C3	C4	C5
DTW Similarity	59.89	54.24	58.31	70.71	72.89
SBFP Similarity	21.41	21.55	29.94	32.05	35.5
Positioning Error	8.62 m	1.46 m	7.93 m	15.56 m	16.34 m

suggest that, the differences among MFS sequences are mainly attributed to the learnt sequence-based fingerprint, which is extracted from small local segments of raw MFS sequence.

We further report the mean positioning error of each test trajectory for the proposed method and baseline positioning models in Table 7. Clearly, the proposed method achieves the best positioning performance on 7 test trajectories, i.e., test path 1,2,3,4,7,8,11, SKNN and SDTW achieve the best positioning performance on two test trajectories, respectively. The results suggest again the superiority of the proposed scale invariant fingerprint for indoor magnetic-based positioning. Our explanation is the learnt scale invariant fingerprint with high discriminative power from local segments of raw MFS sequences can capture the inherent pattern in MFS sequences, thus can boost the positioning performance. On the contrary, the raw MFS sequences may be locally distorted or scaled due to a few factors (e.g., mobile phone pose, metallic materials, ferromagnetic metals and noise in magnetometer measurement) thus greatly limit the positioning performance.

On the other hand, the proposed method does not achieve the best positioning performance on 4 test trajectories, i.e., test path 5, 6, 9, 10. The reason is some training MFS samples that are not ground truth have more segments with the same characteristics with test samples extracted from the four trajectories. For example, for a test sample extracted from trajectory 5 as shown in Figure 8 (1), the proposed method select 5 candidates (i.e., C1, C2, C3, C4 and C5) as shown Figure 8(2)-(6) and related values are shown in Table tab:example. According to SBFP similarity, Candidate C1 is selected as the ultimate result and the positioning error is 8.62 m. But if we use DTW similarity, Candidate C2 is selected as the ultimate result and the positioning error is only 1.46 m. The results suggest that, the proposed sequence-based fingerprint from local patterns does not have the best performance in some scenarios. We plan to improve sequence-based fingerprint by jointly considering some global features (e.g., DTW similarity and curve distance) in the future.

2) EXPERIMENTAL RESULTS OF OUR DATASET

In order to evaluate the performance of the proposed method during different situations, we perform two experiments on our dataset. The first experiment aims to verify the experimental results with different heights, dates and device types. In this experiment, each MFS sequence of training and test dataset is divided in several subsamples with 5 seconds for each one. In this case the training set is made of MFS sequence samples of situation not included in the test set. This evaluation is also known as situation independent evaluation and shows the feasibility of a real positioning application for indoor environment where data of a given situation are usually not included in the training set of the classifier. For example, if the test MFS samples are collected by Xiaomi 5S, the training dataset should only include MFS samples collected by Huawei Mate 10 and Vivo Y87. The second experiment aims to verify both the mean positioning error with different length of a MFS subsequence. In the case, each MFS sequence of training and test dataset is divided in several subsamples from 2 to 10 seconds for each one, correspondingly, the length of MFS sequence vary from 10 to 100.

TABLE 9. Mean positioning error with different data collection dates.

	PKNN	SBOW	SKNN	SDTW	Our method
Date 1	3.27 m	2.76 m	3.01 m	2.63 m	**2.15 m**
Date 2	3.21 m	2.72 m	3.07 m	2.7 m	**2.05 m**
Date 3	3.38 m	2.81 m	3.13 m	2.85 m	**2.21 m**

TABLE 10. Mean positioning error with different data collection devices.

	PKNN	SBOW	SKNN	SDTW	Our method
Xiaomi 5S	3.35 m	2.83 m	3.24 m	2.77 m	**2.18 m**
Huawei Mate 10	3.28 m	2.79 m	3.36 m	2.82 m	**2.21 m**
Vivo Y87	3.48 m	3.01 m	3.56 m	2.95 m	**2.26 m**

TABLE 11. Mean positioning error with different data collection heights.

	PKNN	SBOW	SKNN	SDTW	Our method
Height 1 (0.2 m)	4.07 m	3.75 m	3.93 m	3.72 m	**2.34 m**
Height 2 (1 m)	3.75 m	3.28 m	3.49 m	3.31 m	**2.2 m**
Height 3 (1.8 m)	3.81 m	3.22 m	3.51 m	3.35 m	**2.26 m**

Table 9,10 and 11 show the mean positioning error of the proposed positioning method compare with four baseline methods during different situations. From the three tables, we observe: 1) the proposed method performs much better than the other four baseline methods for different situations, showing again the advantage of indoormagnetic-based positioning by learning scale invariant fingerprint from MFS sequence's local segments. For example, the mean positioning error of our proposed method is 2.05 m when generating test dataset using the MFS sequences collected on date 2,

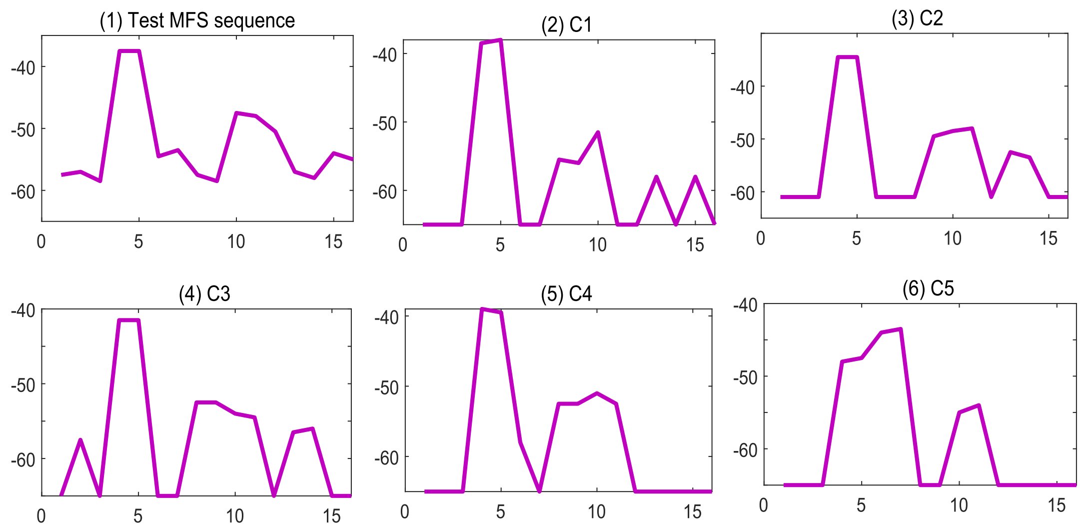

FIGURE 8. A test sample extracted from trajectory 5.

TABLE 12. Mean positioning error (m) with different MFS sequence lengths.

Method \ MFS length	10	20	30	40	50	60	70	80	90	100
SBOW	3.78	3.25	2.99	2.83	**2.75**	2.79	2.93	3.14	3.36	3.72
SKNN	4.03	3.69	3.43	3.27	**3.05**	3.11	3.23	3.48	3.72	3.9
SDTW	3.84	3.3	3.08	2.84	2.72	**2.67**	2.75	2.89	3.15	3.46
Our method	3.27	2.89	2.39	2.23	**2.13**	2.19	2.16	2.34	2.58	2.79

which improves 31.7% and 32.6% than SDTW and SBOW; 2) the mean positioning error of test dataset collected by different dates and devices are relatively stable for all positioning methods, while the mean positioning error of four baseline methods on test dataset collected with height 0.2m degrades significantly compares to test dataset collected with the other two heights. For instance, the mean positioning error of SBOW is 3.75 m for test dataset collected with height 0.2m, which drops 14.3% and 16.5% compare with test dataset collected with height 1m and 1.8m. This is no surprising since the patterns of the MFS variation are more altered with low height is due to the presence of ferromagnetic metals, which is also reported in [24] and [13]. On the other hand, the proposed method performs relatively stable with different data collection heights, showing again the advantage of indoor positioning by learning scale invariant fingerprint from MFS sequence's local segments.

Table 12 shows the mean positioning error of sequence-based indoor positioning methods (SBOW, SKNN, SDTW and our method) with different MFS sequence lengths. From this table, we observe that the positioning performance of the three sequence-based indoor positioning methods (SBOW, SKNN and our method) increase gradually when the MFS sequence length increases from 10 to 50 and then drops when the MFS sequence length is greater than 50, the best results are 2.75m for SBOW, 3.05m for SKNN and 2.13m for our method when the MFS sequence length equals to 50. For SDTW, the trend of change is similar but the best results is 2.67m when the MFS sequence length equals to 60. Nevertheless, the proposed method shows the best performance consistently over MFS sequence lengths as it extracts scale invariant fingerprint with high discriminative power from MFS sequence's local segments. Another observation is the mean positioning error degrades significantly when the MFS sequence length is shorter. This is no surprising since there are few MFS values about each MFS sequence, leading the advantage by sequence-based indoor positioning can be ignored. However, the mean positioning error also degrades significantly when the MFS sequence length is relatively large. After analyzing samples with large positioning errors, we find most of these samples are extracted from the intersections between two corridors but not included in the training set. It can be expected that the positioning performance with long MFS sequence can be further improved if we add the samples that extracted from all the intersections between corridors to training data set.

V. CONCLUSION

This paper proposed an indoor magnetic-based positioning model by learning sequence-based fingerprint from raw magnetic field sequences. The proposed method aims to overcome the main bottleneck of existing methods, i.e., the pattern embedded in raw magnetic field sequences might be locally distorted or scaled, which greatly limits the positioning performance. Firstly, we calibrate the raw magnetic field sequences by denosing with a sliding window-based scheme. Then, we learn the sequence-based fingerprint with high discriminative power for each magnetic field sequence by evaluating the prediction quality of numerous candidate magnetic fingerprints. Experimental results on a large-scale benchmark dataset and our collected dataset show that the proposed method achieves much better performance than the state-of-the-art baseline methods for indoor magnetic-based positioning, showing the superiority of our proposed positioning model and also supporting the assumption that the learnt sequence-based fingerprint from raw magnetic field sequences can significantly boost positioning performance.

As future work, we plan to boost WiFi-based fingerprint positioning by the proposed sequence-based fingerprints, and enable real-time indoor magnetic-based positioning by implementing our method using fog computing framework.

REFERENCES

[1] J. She, J. Crowcroft, H. Fu, and F. Li, "Convergence of interactive displays with smart mobile devices for effective advertising: A survey," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 10, no. 2, 2014, Art. no. 17.

[2] M. S. Hossain, "Cloud-supported cyber–physical localization framework for patients monitoring," *IEEE Syst. J.*, vol. 11, no. 1, pp. 118–127, Mar. 2017.

[3] Y. Chen, Z. Zheng, S. Chen, L. Sun, and D. Chen, "Mining customer preference in physical stores from interaction behavior," *IEEE Access*, vol. 5, pp. 17436–17449, 2017.

[4] Y. Chen, J. Zhang, M. Guo, and J. Cao, "Learning user preference from heterogeneous information for store-type recommendation," *IEEE Trans. Services Comput.*, to be published.

[5] Y. Chen, J. Zhang, M. Guo, and J. Cao, "Understanding customer behaviour in urban shopping mall from WiFi logs," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2017, pp. 50–53.

[6] B. Großwindhager, M. Rath, J. Kulmer, S. Hinteregger, M. Bakr, C. A. Boano, K. Witrisal, and K. Römer, "UWB-based single-anchor low-cost indoor localization system," in *Proc. 15th ACM Conf. Embedded Netw. Sensor Syst.*, 2017, Art. no. 34.

[7] J. Li, G. Feng, W. Wei, C. Luo, L. Cheng, H. Wang, H. Song, and Z. Ming, "PSOTrack: A RFID-based system for random moving objects tracking in unconstrained indoor environment," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4632–4641, Dec. 2018.

[8] M. Terán, J. Aranda, H. Carrillo, D. Mendez, and C. Parra, "IoT-based system for indoor location using Bluetooth low energy," in *Proc. IEEE Colombian Conf. Commun. Comput. (COLCOM)*, Aug. 2017, pp. 1–6.

[9] Y. Chen, M. Guo, J. Shen, and J. Cao, "GraphLoc: A graph-based method for indoor subarea localization with zero-configuration," *Pers. Ubiquitous Comput.*, vol. 21, no. 3, pp. 489–505, 2017.

[10] D. Yan, B. Kang, H. Zhong, and R. Wang, "Research on positioning system based on Zigbee communication," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Oct. 2018, pp. 1027–1030.

[11] Y. Shu, K. G. Shin, T. He, and J. Chen, "Last-mile navigation using smartphones," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 512–524.

[12] B. Gozick, K. P. Subbu, R. Dantu, and T. Maeshiro, "Magnetic maps for indoor navigation," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 12, pp. 3883–3891, Dec. 2011.

[13] H. Xie, T. Gu, X. Tao, H. Ye, and J. Lv, "MaLoc: A practical magnetic fingerprinting approach to indoor localization using smartphones," in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput.*, 2014, pp. 243–253.

[14] Z. Meng, M. Wang, E. Wang, and X. Xu, "Robust local weighted regression for magnetic map-based localization on smartphone platform," *J. Comput. Commun.*, vol. 5, no. 3, p. 80, 2017.

[15] V. Pasku, A. De Angelis, M. Dionigi, G. De Angelis, A. Moschitta, and P. Carbone, "A positioning system based on low-frequency magnetic fields," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2457–2468, Apr. 2016.

[16] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, and M. Wiseman, "Indoor location sensing using geo-magnetism," in *Proc. 9th Int. Conf. Mobile Syst., Appl., Services*, 2011, pp. 141–154.

[17] E. Le Grand and S. Thrun, "3-axis magnetic field mapping and fusion for indoor localization," in *Proc. IEEE Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Sep. 2012, pp. 358–364.

[18] C. E. Galván-Tejada, J. P. García-Vázquez, and R. F. Brena, "Magnetic field feature extraction and selection for indoor location estimation," *Sensors*, vol. 14, no. 6, pp. 11001–11015, 2014.

[19] M. Frassl, M. Angermann, M. Lichtenstern, P. Robertson, B. J. Julian, and M. Doniec, "Magnetic maps of indoor environments for precise localization of legged and non-legged locomotion," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2013, pp. 913–920.

[20] Q. Wang, H. Luo, F. Zhao, and W. Shao, "An indoor self-localization algorithm using the calibration of the online magnetic fingerprints and indoor landmarks," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2016, pp. 1–8.

[21] N. Lee, S. Ahn, and D. Han, "AMID: Accurate magnetic indoor localization using deep learning," *Sensors*, vol. 18, no. 5, p. 1598, 2018.

[22] Z. Liu, L. Zhang, Q. Liu, Y. Yin, L. Cheng, and R. Zimmermann, "Fusion of magnetic and visual sensors for indoor localization: Infrastructure-free and more effective," *IEEE Trans. Multimedia*, vol. 19, no. 4, pp. 874–888, Apr. 2017.

[23] A. Solin, S. Särkkä, J. Kannala, and E. Rahtu, "Terrain navigation in the magnetic landscape: Particle filtering for indoor positioning," in *Proc. Eur. Navigat. Conf. (ENC)*, May/Jun. 2016, pp. 1–9.

[24] H. Xie, T. Gu, X. Tao, H. Ye, and J. Lu, "A reliability-augmented particle filter for magnetic fingerprinting based indoor localization on smartphone," *IEEE Trans. Mobile Comput.*, vol. 15, no. 8, pp. 1877–1892, Aug. 2016.

[25] R. Ban, K. Kaji, K. Hiroi, and N. Kawaguchi, "Indoor positioning method integrating pedestrian dead reckoning with magnetic field and WiFi fingerprints," in *Proc. 8th Int. Conf. Mobile Comput. Ubiquitous Netw. (ICMU)*, Jan. 2015, pp. 167–172.

[26] Y. Li, Y. Zhuang, H. Lan, Q. Zhou, X. Niu, and N. El-Sheimy, "A hybrid WiFi/magnetic matching/PDR approach for indoor navigation with smartphone sensors," *IEEE Commun. Lett.*, vol. 20, no. 1, pp. 169–172, Jan. 2016.

[27] X. Huang, S. Guo, Y. Wu, and Y. Yang, "A fine-grained indoor fingerprinting localization based on magnetic field strength and channel state information," *Pervasive Mobile Comput.*, vol. 41, pp. 150–165, Oct. 2017.

[28] Y. Du, T. Arslan, and A. Juri, "Camera-aided region-based magnetic field indoor positioning," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2016, pp. 1–7.

[29] K. P. Subbu, B. Gozick, and R. Dantu, "Locateme: Magnetic-fields-based indoor localization using smartphones," *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 4, p. 73, 2013.

[30] R. Montoliu, J. Torres-Sospedra, and O. Belmonte, "Magnetic field based indoor positioning using the bag of words paradigm," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2016, pp. 1–7.

[31] M. Angermann, M. Frassl, M. Doniec, B. J. Julian, and P. Robertson, "Characterization of the indoor magnetic field for applications in localization and mapping," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Nov. 2012, pp. 1–9.

[32] J. Song, H. Jeong, S. Hur, and Y. Park, "Improved indoor position estimation algorithm based on geo-magnetism intensity," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2014, pp. 741–744.

[33] W. Storms, J. Shockley, and J. Raquet, "Magnetic field navigation in an indoor environment," in *Proc. Ubiquitous Positioning Indoor Navigat. Location Based Service (UPINLBS)*, Oct. 2010, pp. 1–10.

[34] J. Torres-Sospedra, D. Rambla, R. Montoliu, O. Belmonte, and J. Huerta, "Ujiindoorloc-mag: A new database for magnetic field-based localization problems," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2015, pp. 1–10.

[35] C.-C. M. Yeh, Y. Zhu, L. Ulanova, N. Begum, Y. Ding, H. A. Dau, Z. Zimmerman, D. F. Silva, A. Mueen, and E. Keogh, "Time series joins, motifs, discords and shapelets: A unifying view that exploits the matrix profile," *Data Mining Knowl. Discovery*, vol. 32, no. 1, pp. 83–123, 2018.

[36] J. Hills, J. Lines, E. Baranauskas, J. Mapp, and A. Bagnall, "Classification of time series by shapelet transformation," *Data Mining Knowl. Discovery*, vol. 28, no. 4, pp. 851–881, 2014.

[37] J. Lines, L. M. Davis, J. Hills, and A. Bagnall, "A shapelet transform for time series classification," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 289–297.

YUANYI CHEN received the B.Sc. degree from Sichuan University, in 2010, the M.Sc. degree from Zhejiang University, in 2013, and the Ph.D. degree from Shanghai Jiao Tong University, in 2017. He was a Visiting Scholar under the supervision of Prof. K.-J. Lin with the University of California at Irvine, from November 2018 to November 2019. He is currently a Distinguished Research Fellow with the Department of Computer Science and Computing, Zhejiang University City College. His research interests include the Internet of Things, mobile computing, and ubiquitous computing. He has published more than 20 technical articles in major international journals and conference proceedings.

MINGXUAN ZHOU received the B.S. degree in electrical engineering from Zhejiang University, in 2017. He is currently pursuing the master's degree with the College of Computer Science and Technology, Zhejiang University, and the Zhejiang University City College. His research topic includes the Internet of Things and ubiquitous computing.

ZENGWEI ZHENG received the B.S and M.Ec. degrees in computer science and western economics from Hangzhou University, China, in 1991 and 1998, respectively, and the Ph.D. degree in computer science and technology from Zhejiang University, China, in 2005. He is currently a Professor with the Department of Computer Science and Engineering, the Director of the Intelligent Plant Factory of Zhejiang Province Engineering Lab and the Hangzhou Key Laboratory for the IoT Technology and Application, and the Head of the Department of Scientific Research, Zhejiang University City College. His research interests include wireless sensor networks, location-based service, the Internet of Things, digital agriculture, and pervasive computing. He is also a member of the ACM and the CCF.