

# Self-Attention-Enhanced PSW-LSTM for 3-D Indoor Pedestrian Positioning With Integrated Wi-Fi, Magnetometer, and Barometer Sensors

Zhanpeng Zhang<sup>ID</sup>, Jiale Wang<sup>ID</sup>, Member, IEEE, Ming Xia<sup>ID</sup>, Chuang Shi<sup>ID</sup>, Weisong Wen<sup>ID</sup>, Member, IEEE, and Deyou Zhang<sup>ID</sup>, Member, IEEE

**Abstract**—The performance of smartphone-based fingerprinting indoor positioning methods heavily depends on the database quality and matching algorithms, typically providing only 2-D coordinates. Traditional methods also face challenges such as labor-intensive data collection, limited accuracy, and reduced long-term reliability. This article proposes a self-attention-enhanced presequence warmup long short-term memory (PSW-LSTM) framework, integrated with Wi-Fi, barometer, and magnetometer sensors, to achieve accurate 3-D pedestrian indoor positioning. Specifically, foot-mounted positioning devices are introduced to assist smartphones in collecting sparse fingerprint data, improving database efficiency. The proposed self-attention-enhanced PSW-LSTM framework improves initial matching accuracy, reduces convergence time, and enhances trajectory continuity. Experimental results demonstrate that, in typical indoor environments such as educational and office spaces, the horizontal and vertical positioning accuracy for pedestrians is better than 2 and 0.5 m, significantly improving smartphone-based 3-D indoor positioning performance.

**Index Terms**—3-D indoor positioning, multisource sensors fusion, rapid fingerprint database establishment, self-attention-enhanced presequence warmup long short-term memory (PSW-LSTM), smartphone.

## I. INTRODUCTION

**G**NSS, as a significant positioning method, is a global hotspot in satellite navigation applications and technology research. However, its effectiveness is hindered by various

Received 11 November 2024; revised 25 March 2025; accepted 18 April 2025. Date of publication 12 May 2025; date of current version 28 May 2025. This work was supported in part by the Science and Technology Project of the State Grid Corporation of China (Research and Application of Key Technologies for Autonomous and Controllable Power Beidou High-Reliability Space-Time Service) under Grant 5700-202441237A-1-ZN, in part by Beihang Ganwei Program under Grant WZ2024-2-31, and in part by Beihang World TOP University Cooperation Program. The Associate Editor coordinating the review process was Dr. Gui-Bin Bian. (*Corresponding authors: Ming Xia; Jiale Wang.*)

Zhanpeng Zhang and Deyou Zhang are with the School of Electronics and Information Engineering, Beihang University, Beijing 100191, China (e-mail: zhanpeng\_zhang@buaa.edu.cn; deyou@buaa.edu.cn).

Jiale Wang and Chuang Shi are with the School of Space and Earth Sciences, Beihang University, Beijing 100191, China, and also with the Key Laboratory of Satellite Navigation and Mobile Communication Fusion Technology, Ministry of Industry and Information Technology, and the National Key Laboratory of CNS/ATM, Beijing 100191, China (e-mail: wang\_jiale@buaa.edu.cn; shichuang@buaa.edu.cn).

Ming Xia is with the School of Electronics and Information Engineering, Beihang University, Beijing 100191, China, and also with Jiangxi Research Institute, Beihang University, Nanchang, Jiangxi 330006, China (e-mail: xiaming@buaa.edu.cn).

Weisong Wen is with the Department of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University, Hong Kong 999077, China (e-mail: welson.wen@polyu.edu.hk).

Digital Object Identifier 10.1109/TIM.2025.3568977

factors, making it insufficient to meet indoor positioning needs solely with GNSS [1], [2]. Therefore, achieving “last kilometer” indoor positioning has become an urgent challenge to address. Currently, there are many commonly used indoor positioning methods, such as the inertial navigation system [3], [4], [5], visual simultaneous localization and mapping (SLAM) [6], [7], [8], map matching [9], [10], acoustic positioning [11], ultrawideband (UWB) [12], and so on [13], [14], [15], [16].

Using existing signals in the environment for indoor positioning is also a common approach. This typically involves leveraging wireless networks (such as Wi-Fi [17], [18] and Bluetooth [19], [20]) or inherent characteristics of the Earth (such as geomagnetism [21], [22] and barometric pressure [23]) for positioning. This approach typically involves two stages: offline data collection and training, followed by online matching and positioning. In the first stage, several reference points are preselected in the target environment, and signal measurements are taken at each reference point. The signal characteristics of each reference point are recorded to construct a signal fingerprint database. In the subsequent positioning phase, the signal characteristics of the target point to be located are collected again and matched with the signals in the fingerprint database. It usually does not require additional hardware, thereby reducing deployment costs. Additionally, in complex environments with severe multipath effects and signal attenuation, it can capture environmental characteristics through a large amount of sample data, thereby achieving more stable positioning.

However, traditionally, the construction of a fingerprint database incurs high costs, requiring extensive signal collection efforts in the area. This process not only necessitates the use of relatively precise instruments but is also time-consuming and labor-intensive. Additionally, during the positioning phase, significant signal-matching computations are required, which places high demands on computational resources and response speed. Furthermore, in areas where signal strength varies greatly or signal characteristics are not distinct, positioning accuracy may decrease.

In recent years, positioning methods utilizing signal fingerprints have undergone continuous optimization and improvement. Regarding the construction of fingerprint databases, Tong et al. [24] proposed a novel approach leveraging Wi-Fi channel state information (CSI) for high-precision indoor localization. Their method utilizes CSI phase difference as fingerprints and employs mini-computers for data

collection, to improve the efficiency of database construction. However, this approach still relies on specialized hardware devices for data acquisition, limiting its universal applicability. Lu et al. [25] introduced ONavi, a system that utilizes smartphone sensors, particularly Wi-Fi and IMU, for data acquisition. This approach enhances accessibility and practicality, as it does not require specialized hardware. However, while leveraging smartphone sensors is advantageous for widespread deployment, the relatively lower precision of smartphone IMUs may potentially compromise the reliability of the fingerprint database during the construction phase. Tan et al. [26] presented UbiFin, a system that employs implicit crowdsourcing of multimodal signals including RF, geomagnetic fields, and motion data. This crowdsourcing approach enhances efficiency and scalability in fingerprint collection, as it does not require explicit user participation. However, it also presents challenges such as uneven data coverage across different areas and inconsistent data quality due to variations in user devices and behaviors. Therefore, balancing the convenience of the device with the precision of data measurement, while ensuring high-efficiency data collection and maintaining high data quality, remains a critical issue that requires further exploration and discussion. Moreover, the majority of existing research has primarily focused on planar spaces or floor identification, as exemplified by Gu et al.'s [27] FloorLocator. These studies often limit their scope to common hallways or single rooms, failing to address the implementation of multiscene fingerprint localization in 3-D spaces. This narrow focus leaves a significant gap in our understanding and capability to perform comprehensive indoor positioning across diverse spatial contexts.

Furthermore, once fingerprint data is acquired, appropriate matching algorithms are essential to achieve comprehensive localization. Zhu et al. [28] proposed an indoor positioning algorithm that innovatively employs a Rao–Blackwellized particle filter to divide the environment into local areas, effectively reducing computational complexity while maintaining accuracy, yet its performance in more complex, large-scale environments remains to be fully validated. With the rise of machine learning methods, many researchers have integrated these techniques into indoor positioning systems, further enhancing the performance. Zhao et al. [29] addressed the issue of localization method switching by employing a BP neural network and a BP neural network optimized by particle swarm optimization. Chen et al. [30] employed partial least-squares (PLS) to extract features from Wi-Fi fingerprints, and combined with a relevance vector machine (RVM-PLS) for regression and prediction, aiming to enhance localization accuracy. Chen utilized magnetic field data and dynamic time warping (DTW) [31], while Zhang employed Wi-Fi signals and  $k$ -nearest neighbors (KNNs) [32] to achieve high-precision indoor localization, enhancing the system's accuracy and robustness in dynamic and complex environments. Ji et al. [33] utilized the power of 5G signals as fingerprints and applied random forest for variable selection of multivariate fingerprints. They also compared the performance of various machine learning algorithms, including multilayer perceptron. Zhang et al. [34] utilized magnetic field fingerprint data

collected from smartphones in conjunction with long short-term memory (LSTM) networks and validated the method's accuracy, response time, and robustness in localization. However, these algorithms still have considerable room for improvement in terms of positioning accuracy and robustness. Moreover, their applicability remains relatively limited. The effective integration of multisource information continues to pose significant challenges. Additionally, these algorithms are typically tested only with real-time collected data. When a substantial time interval elapses, the fingerprint data may undergo dynamic changes, potentially resulting in a marked degradation of positioning performance. This temporal instability highlights the need for more adaptive and resilient positioning techniques that can maintain accuracy over extended periods in dynamic environments.

To address the issues of low efficiency in fingerprint database construction and insufficient accuracy of traditional matching positioning algorithms, a method combining foot-mounted positioning terminals and smartphones for fingerprint database construction is proposed. Additionally, deep learning techniques are utilized to further enhance positioning accuracy and robustness. This article makes the following key contributions.

- 1) Introduce an improved deep learning framework called self-attention-enhanced PSW-LSTM, designed to significantly enhance both the efficiency and performance of positioning systems. By leveraging the self-attention mechanism and the PSW method, this framework accelerates model convergence, allowing for faster and more effective initialization of matching trajectories and further boosting the accuracy and robustness of positioning, especially in dynamic and complex environments. The proposed method outperforms traditional machine learning approaches by over 50% and surpasses conventional LSTM techniques by approximately 30% in overall localization performance.
- 2) We propose a feature fusion framework that leverages the collaborative capabilities of foot-mounted positioning terminals and smartphones for sparse data collection. The framework enables efficient and rapid construction of a signal fingerprint database, maximizing the strengths of each device to improve the overall quality and coverage of the fingerprint database. Compared to traditional methods using a combination of total stations and multiple sensors, this approach demonstrates an efficiency improvement of at least 20-fold.

This article is structured as follows. Section II presents the proposed self-attention-enhanced PSW-LSTM algorithm. Section III introduces the method for rapidly constructing a signal fingerprint database and the details of the experimental setup. Section IV analyzes the experimental results. Section V provides a summary of the paper and discusses future prospects. Fig. 1 illustrates the framework for constructing the entire system.

## II. DEEP LEARNING FOR MATCHING LOCALIZATION

The core challenge in matching-based positioning is to efficiently identify the signals in the fingerprint database

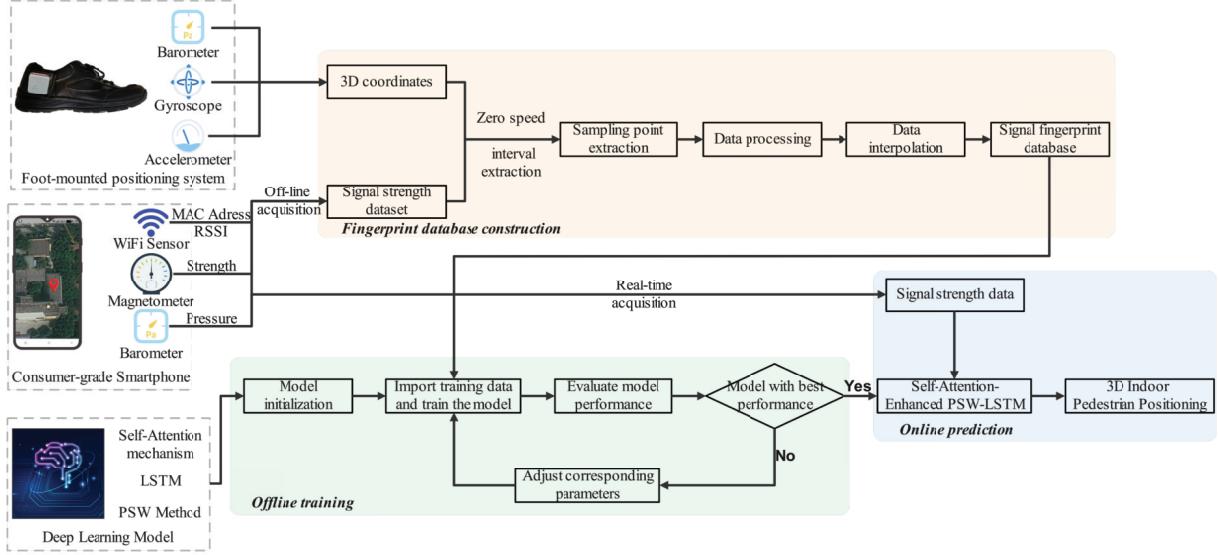


Fig. 1. Framework for constructing the system.

that most closely align with newly collected signal features and to estimate the corresponding location [35]. Due to the potential for multiple data points exhibiting similar signal characteristics, distinguishing between them accurately can be difficult. This may result in discontinuities in the predicted trajectory and increased positioning errors. To mitigate this issue, we propose a model based on LSTM architecture, which effectively addresses these discontinuities and ensures continuous trajectory estimation. Additionally, we introduce specific optimizations to enhance positioning accuracy and robustness, leading to the development of the PSW-LSTM model with self-attention. This model uses multiple signal strengths, such as Wi-Fi, geomagnetic, and barometric signals, as inputs, with the output being the estimated 3-D position coordinates.

#### A. Self-Attention-Enhanced LSTM Network Model

The LSTM network is a specialized type of recurrent neural network (RNN) designed to process and retain long-term sequence data [36]. Unlike traditional RNNs, LSTMs address common challenges such as vanishing and exploding gradients. The core of the LSTM architecture consists of memory cells and three main gates: the forget gate, the input gate, and the output gate. The specific correspondence is illustrated in Fig. 2 and described in the following equation:

$$\begin{cases} f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ g_t = \tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \\ o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\ C_t = C_{t-1} \odot f_t + g_t \odot i_t \\ h_t = o_t \odot \tanh(C_t) \end{cases} \quad (1)$$

where the forget gate  $f_t$ , controlled by a sigmoid function, filters past information, while the input gate  $i_t$  determines whether a new candidate state,  $g_t$ , should be added. The

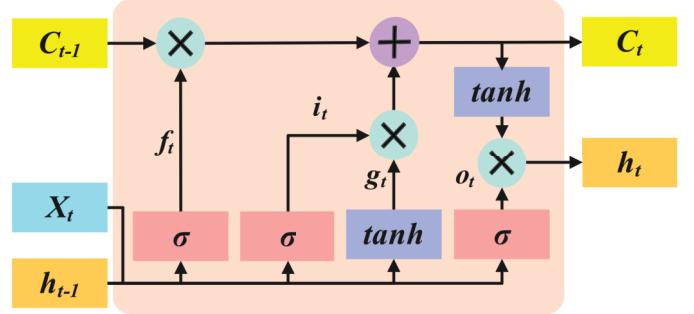


Fig. 2. LSTM network architecture.

output gate  $o_t$ , also controlled by a sigmoid function, selects which memory content influences the current hidden state  $h_t$ .  $W_{xf}$ ,  $W_{xi}$ ,  $W_{xg}$ , and  $W_{xo}$  represent the weight matrices for the inputs to each gate, while  $W_{hf}$ ,  $W_{hi}$ ,  $W_{hg}$ , and  $W_{ho}$  denote the weight matrices for the hidden states to each gate. The bias terms  $b_f$ ,  $b_i$ ,  $b_g$ , and  $b_o$  are introduced to optimize these gates, enabling the model to dynamically manage long-term dependencies in complex spatiotemporal tasks, allowing the network to adaptively adjust and enhance the learning process.

For the LSTM layer of this system, the input  $x_t$  consists of signal strength information, including Wi-Fi, magnetic field, and pressure data, at different time steps and locations. The output  $h_t$  represents the hidden state at various locations. After passing through a Dropout layer, some elements of  $h_t$  are removed, and the remaining elements are passed to the subsequent layers. Finally, a regression layer is applied to output the corresponding 3-D position.

Although LSTM layers have advantages in handling long sequence problems, challenges remain. Despite the gating mechanisms that mitigate long-term dependency issues, as the sequence length increases, early input information may still be gradually diluted. For instance, during prolonged usage, the initial position information may lose its effectiveness in reference. Moreover, while the LSTM hidden states recursively

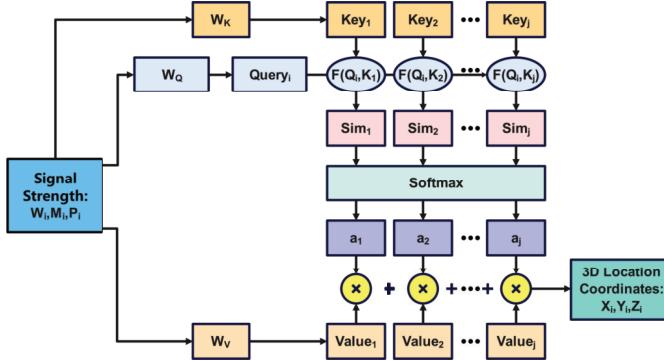


Fig. 3. Design of the self-attention mechanism.

integrate historical information, the contributions from all time steps are implicit and treated equally, making it difficult to explicitly distinguish crucial moments [37]. In contrast, self-attention enables the model to directly compute the association weights between any two time steps, regardless of their distance. This mechanism effectively adjusts the contribution of each dimension dynamically, suppressing transient noise, enhancing relevant signals, and preventing trajectory jumps in localization. The goal of the self-attention mechanism is to assign different weights to parts of the input sequence, enabling better information extraction and processing, as shown in Fig. 3.

For the self-attention mechanism, the input is the hidden state sequence  $S_i$  from the LSTM layer. The weight matrices  $W_K$ ,  $W_Q$ , and  $W_V$  are trained and subsequently used to generate the query ( $Q_i$ ) for target feature querying, the key ( $K_j$ ) for feature matching, and the value ( $V_j$ ) that carries the actual information. The detailed generation process is shown in the following equation:

$$\begin{cases} K_j &= \mathbf{W}_K \odot S_i \\ Q_i &= \mathbf{W}_Q \odot S_i \\ V_j &= \mathbf{W}_V \odot S_i. \end{cases} \quad (2)$$

Subsequently, the correlation between the query and the key is computed to obtain the similarity score  $\text{Sim}_i$ , which is used to quantify the correlation or match between  $Q_i$  and each  $K_j$ . We use the dot product result, as shown in (3), to represent the interaction between the two vectors

$$\text{Sim}_i(Q_i, K_j) = Q_i \cdot K_j. \quad (3)$$

Next, the Softmax function, as depicted in (4), is applied to the similarity scores ( $\text{Sim}_i$ ) to normalize them, yielding the corresponding weight coefficients which are denoted as  $a_i$

$$a_i = \text{Softmax}(\text{Sim}_i) = \frac{e^{\text{Sim}_i}}{\sum_{j=1}^n e^{\text{Sim}_j}}. \quad (4)$$

Finally, the weight coefficients are used to compute a weighted sum of the value vectors, denoted as  $V_i$ , described by (5), resulting in the final output of the self-attention mechanism, a feature sequence based on global context fusion after fusion, denoted as  $L_i$

$$L_i = \text{Attention}(\text{Query}) = \sum_{i=1}^n a_i \cdot V_i. \quad (5)$$

Given that the positioning system operates in a 3-D environment, various scenarios such as staircases are inherently part of the data. In scenarios involving staircases, the self-attention mechanism assigns higher weights to barometric pressure features, as pressure changes are more pronounced during vertical displacement, allowing for more accurate vertical positioning. In contrast, when the user is moving along flat surfaces such as hallways, the system shifts its focus to Wi-Fi and geomagnetic features, which are more reliable for horizontal localization. This dynamic weighting based on the environment ensures that the system effectively captures the relevant information for precise indoor positioning, improving the accuracy of both vertical and horizontal localization across diverse settings.

#### B. Complete PSW-Assisted Network for Matching and Localization

In LSTM models, the reliance on past time steps for current predictions presents a challenge at the beginning of positioning sequences where historical information is insufficient [38].

This lack of sufficient prior data can cause the model's initial hidden and cell states to be inadequately adjusted, leading to difficulties in early predictions and potential drift from the starting point. To address this issue, we introduce the presequence warmup-LSTM (PSW-LSTM) method. Before formal testing, sequences unrelated to the test set are used to ensure that the model reaches a stable state. These warmup sequences, collected near the starting point before the test data is acquired, are not used for training or testing to prevent data leakage. Predictions made during this phase are excluded from the final positioning performance metrics. Although the warmup data comes from the same environment and conditions as the test data, its sole purpose is to ensure that the model's initial state closely matches the actual testing conditions. By stabilizing the model's initial state, this approach mitigates cold start issues and reduces initial point drift, ultimately improving positioning accuracy.

Fig. 4 presents the complete system model. The LSTM layer processes sensor data (such as Wi-Fi, magnetic field, and pressure readings) to capture the user's movement over time. The output represents abstract features derived from the data, not raw measurements or final location estimates. A dropout layer is applied to randomly remove some of these features during training, which helps prevent overfitting and improves the model's ability to generalize. Next, fully connected layers combine these features to produce the final output, which in this case is a t3-D location estimate. Essentially, the model learns to identify the most relevant aspects of the sensor data to determine the user's indoor position, ensuring accurate and reliable localization.

### III. EXPERIMENTAL SETUP

#### A. Construction of the Fingerprint Database

To construct a reliable fingerprint database, three key challenges need to be addressed: accurate acquisition of position coordinates, multisensor signal data collection, and data interpolation to enrich the fingerprint database.

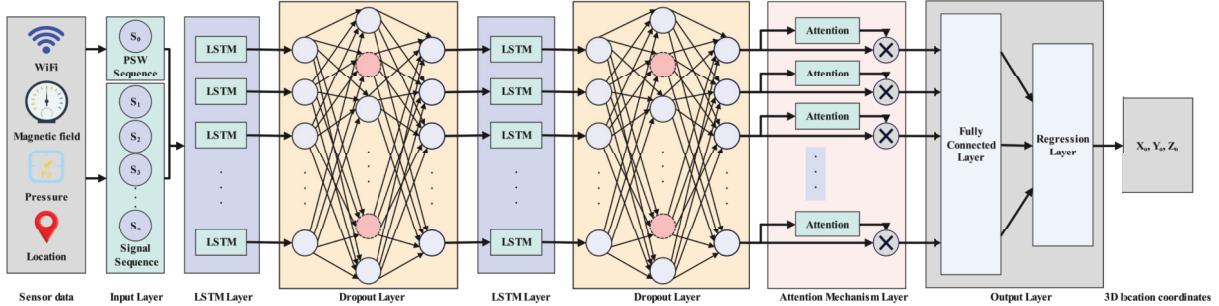


Fig. 4. Self-attention-enhanced PSW-LSTM network framework.

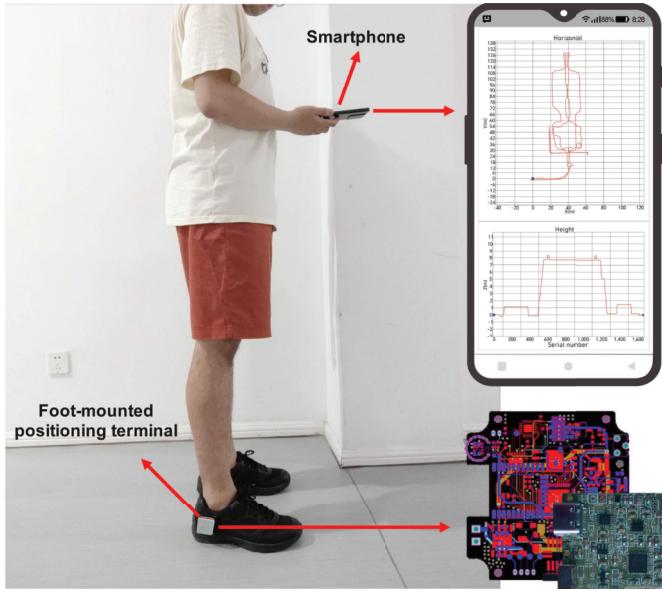


Fig. 5. Experimental equipment setup.

To get the signal fingerprint data, the experiment was conducted by having participants wear a custom-developed positioning terminal [1], [39], relying on the laboratory's platform, on their foot, and carry a Huawei Mate 40 smartphone to collect various signal strengths, as illustrated in Fig. 5.

The first step involves utilizing foot-mounted positioning terminals for the precise collection of position coordinates. This device is essentially a positioning system with a high-precision ten-axis IMU, a communication module, a voltage module, a power supply module, and a high-speed CPU. It is equipped with MEMS-based instruments, to continuously acquire positional data in indoor environments. Moreover, by employing techniques such as adaptive zero-velocity updates (AZUPT), pattern recognition (including modes such as standing still, walking, climbing stairs, and elevator usage), and sensor-adaptive calibration, the system effectively mitigates the cumulative errors inherent in inertial navigation. Consequently, the horizontal trajectory closure error is maintained below 0.2%. This system has been successfully implemented in various large-scale companies such as the State Grid Southern China and Guangdong Fire Department [1], [40]. In terms of vertical positioning accuracy, the foot terminal was able to accurately identify floors. To ensure the accuracy of the

reference coordinates provided, the height difference between floors as well as the height of each stair step was precalibrated using a handheld laser rangefinder. By integrating this precalibrated height information with the foot-pedal inertial navigation system, we obtained accurate reference coordinates for the training and test sets. Additionally, the terminal's sensors record coordinates at a frequency of up to 1000 Hz, while a typical pedestrian takes approximately 0.5 s per step, meaning at least 500 coordinate data points can be recorded per step. As a result, the relevant coordinates are synchronously logged with the user's movement, eliminating the need for additional waiting time. In contrast, traditional methods using a total station for coordinate collection require 2–5 minutes per point, depending on environmental conditions and the need for instrument calibration and initialization. For an indoor environment with 100 measurement points, using a total station would take at least 200 min, whereas the foot-mounted terminal would complete the collection in under 1 min, leading to an overall efficiency improvement of over 200 times, drastically reducing the time required for coordinate acquisition. Furthermore, the terminal is also capable of recording the zero-velocity information of the device. When the device is in a zero-velocity state (stationary) or a non-zero-velocity state (walking), it is recorded as 1 and 0, respectively. This allows for the filtering of valid data points corresponding to the zero-velocity state (1) during postprocessing, which enhances the quality of the database better. Ultimately, the position and zero-velocity information collected using this terminal will be recorded as  $x_i$ ,  $y_i$ ,  $z_i$ , and  $\text{zupt}_i$  for every point.

For Wi-Fi, magnetic field, and pressure data, traditional approaches typically rely on separate sensors for individual measurements, with the data later aggregated. However, smartphones are equipped with all these sensors, allowing for more streamlined data collection. Through a laboratory-developed app, these sensors can be effectively utilized to capture and integrate signal data, enabling a more efficient and cohesive measurement process [41].

The software will transmit the position data from the foot-mounted terminal to the smartphone via Bluetooth, displaying both horizontal and vertical walking trajectories and their variations. Additionally, upon start-up, the software will automatically activate the Wi-Fi module, magnetometer, barometer, and other sensors and begin recording signal strength information corresponding to each position. The Wi-Fi signals capture

the media access control (MAC) addresses of detected networks along with their corresponding received signal strength (RSS). While device heterogeneity may introduce variability in low-RSS measurements, high-RSS values (primarily employed for fingerprint matching) exhibit reduced susceptibility to such discrepancies, as weak signals are inherently assigned lower weights in localization algorithms. Consequently, RSS variations induced by device orientation differences are deemed negligible in practical implementations.

For barometric pressure data, relative pressure values are systematically utilized rather than absolute measurements to mitigate device-specific calibration biases. Specifically, the starting point, located on the same floor, is used as a reference point. By recording the pressure changes relative to this reference point, the system effectively captures the variations as relative pressure. This relative pressure is treated as the pressure fingerprint, providing a more stable and consistent metric for indoor positioning. Using this approach, we minimize environmental factors and calibration discrepancies, ensuring that the pressure data remains reliable throughout the experiment [42], [43].

Regarding geomagnetic field intensity, raw sensor measurements acquired directly from mobile devices inherently reside in the device's local coordinate system, rendering them orientation-dependent if used without transformation. To address this, we implement a sensor fusion framework leveraging accelerometer and gyroscope data to project magnetic field vectors into the standardized east-north-up (ENU) navigation coordinate system, thereby decoupling positioning accuracy from arbitrary device orientations.

Additionally, it is important to note that the signal acquisition frequencies vary among different sensors. To ensure data integrity, we will remain stationary for 5 s at each sampling point. Even with the pause at each sampling point, the collection of 100 data points only required 500 s. Compared to traditional methods using a combination of total stations and multiple sensors, this approach demonstrates an efficiency improvement of at least 20-fold! Subsequently, we will extract the zero-velocity information from the foot-mounted terminal during postprocessing. The average values over this period will be calculated to obtain comprehensive fingerprint data with higher reliability. Finally, at a certain point, the Wi-Fi, magnetic field, and air pressure values recorded by the smartphone will be denoted as  $W_i$ ,  $M_i$ , and  $P_i$  respectively. Here,  $W_i$  includes the Wi-Fi signal strengths for all different MAC addresses and  $M_i$  includes the magnetic field strengths in the north, east, and vertical directions.

To further enhance the efficiency of fingerprint database construction, this article employs a method of sparse fingerprint collection combined with interpolation. The interpolation technique utilized is called dual-side inverse distance weighting (DS-IDW). Specifically, an equal number of sampling points are selected from both the left and right sides of the point to be interpolated as reference points. The signal strength is then estimated using the distance relationship between the point to be interpolated and the reference points as weights. The specific theoretical representation is shown in

TABLE I  
SPECIFIC INTERPOLATION ERROR

Magnetic field component	RMSE ( $\mu T$ )	Max ( $\mu T$ )
East	0.2	0.6
North	2.6	9.2
Up	3.3	10.0

the following equation:

$$D_n(X_n, Y_n) = \frac{\sum_{i=1}^n \frac{D(X_i, Y_i)}{(d_i)^P}}{\sum_{i=1}^n \frac{1}{(d_i)^P}} \quad (6)$$

where  $D_n(X_n, Y_n)$  is the signal strength data interpolated at the point  $(X_n, Y_n)$ ,  $D_i(X_i, Y_i)$  is the signal strength data at the reference point  $(X_i, Y_i)$ ,  $d_i$  is the distance between the point to be interpolated and the reference point,  $P$  is the weight coefficient, and  $n$  is the total number of reference points selected.

To verify the accuracy of the values obtained by this method, an interpolation comparison was conducted within a certain area. Fig. 6 shows a comparison between a measured real magnetic field sequence and the magnetic field sequence obtained through interpolation. The blue curve represents the real data (reference sequence) and the red trajectory represents the interpolated data (interpolated sequence). The shape and trend of the sequences are basically the same, with only some differences in numerical values. The specific statistical results are shown in Table I. The overall rms values are better than  $3.5 \mu T$ , and the overall interpolation accuracy is within an acceptable range.

The signal fingerprint database, along with the testing data, has been made openly accessible to the public ([https://github.com/PengBUAA/Dataset\\_IEEE\\_TIM](https://github.com/PengBUAA/Dataset_IEEE_TIM)). The test set will be continuously updated with new data, enabling ongoing experimentation, validation, and refinement of positioning technologies.

### B. Experiment in the East Annex of the Library

To verify the accuracy of the positioning, two sets of experiments were designed in this study. The first experiment is conducted on the first and second floors of the East Annex of the Library at Beihang University. Fig. 7 presents the satellite image, 3-D schematic, and the corresponding walking trajectories for the building. The yellow trajectory represents the path formed by connecting the reference points in the training set, while the cyan trajectory represents the path formed by connecting the points in the testing set.

This experiment not only serves to validate the localization performance of the algorithm across diverse scenarios but also addresses a crucial aspect of fingerprint-based positioning. It is noteworthy that fingerprint information typically undergoes dynamic changes over time. These changes are typically caused by environmental factors such as variations in temperature and humidity, human activity and density, and alterations to the indoor structure. The signal fingerprint variations can exceed 2–15 dB. If these changes are not adequately addressed, they can significantly impact localization

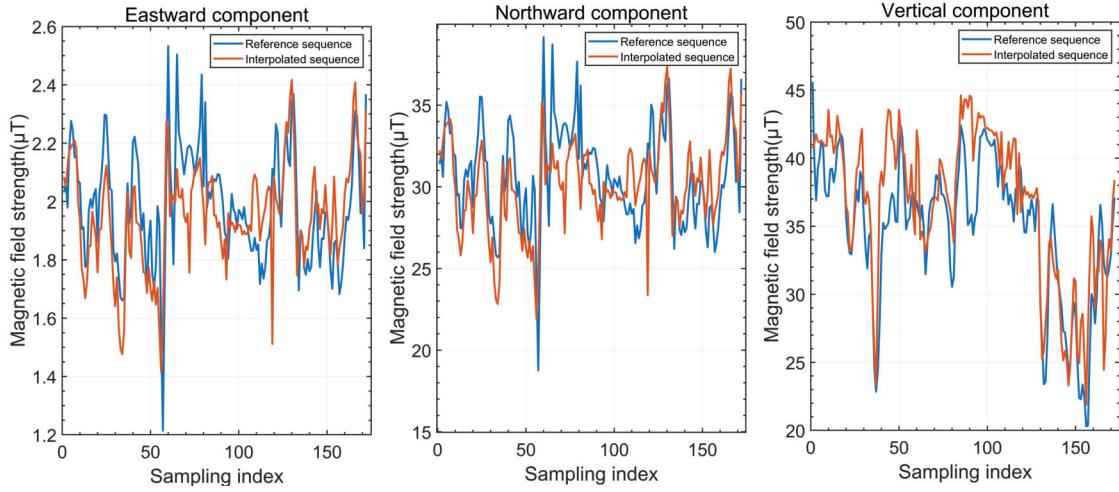


Fig. 6. Comparison between reference data and interpolated data.

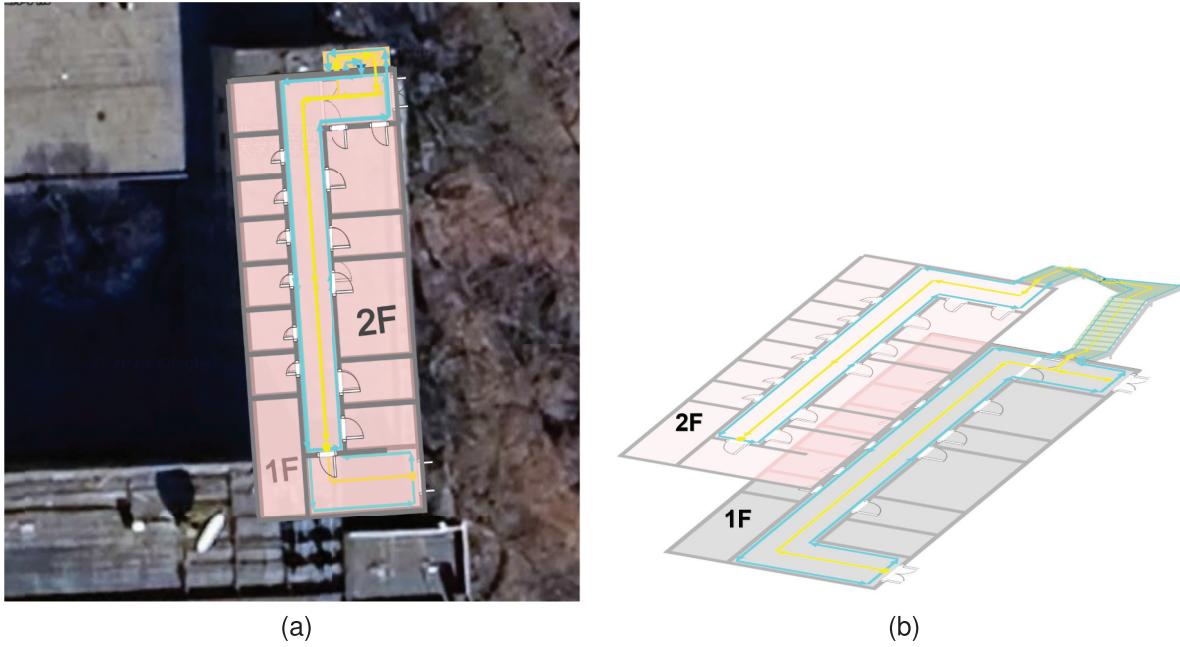


Fig. 7. Satellite image and 3-D schematic of the East Annex of the Library. (a) Satellite image and (b) 3-D schematic.

accuracy. Therefore, it is essential to comprehensively evaluate the performance of the method both in real time and over extended periods to ensure long-term stability and reliability. In practice, the frequency of updating the fingerprint database depends on the dynamics of the environment. For example, in environments with frequent changes, such as areas with high human traffic, the location database typically needs to be updated on a weekly to monthly basis [44], [45], [46]. However, for this experimental environment, university office buildings, although there will be some changes, the overall environment is relatively stable and the fingerprint does not change too frequently. Therefore, to mitigate the effects of such dynamic changes, our method incorporates the warmup calibration process at the beginning of each session. This warmup phase ensures that the model's initial state is aligned with the current environmental conditions, effectively reducing

the need for continuous recalibration during operation. In our experiments, we observed that after performing the warmup process, the localization performance remained stable over the subsequent use period.

Consequently, we have designed a data collection protocol that includes initial collection and subsequent recollection. Specifically, the first data collection phase will involve traversing the yellow trajectory illustrated in Fig. 7, gathering data that will serve as reference values for interpolation and, in conjunction with the interpolated data, constitute the training set. Subsequently, data will be collected along the cyan trajectory to form Test Set 1. The second phase of data collection will be conducted two months after the initial collection. During this phase, only the cyan trajectory will be followed, resulting in the creation of Test Set 2. The localization errors from this second phase will be compared against those from the

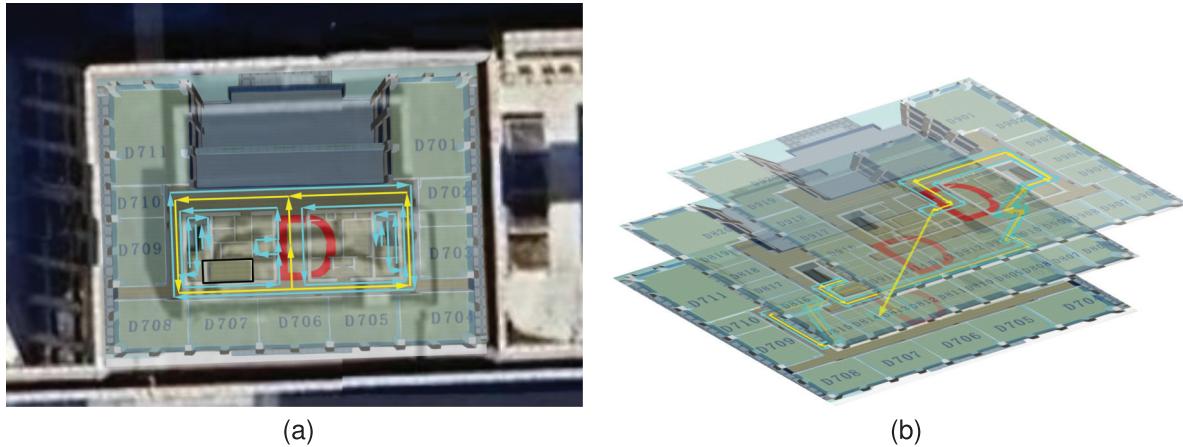


Fig. 8. Satellite image and 3-D schematic of Building D of the New Main Building. (a) Satellite image and (b) 3-D schematic.

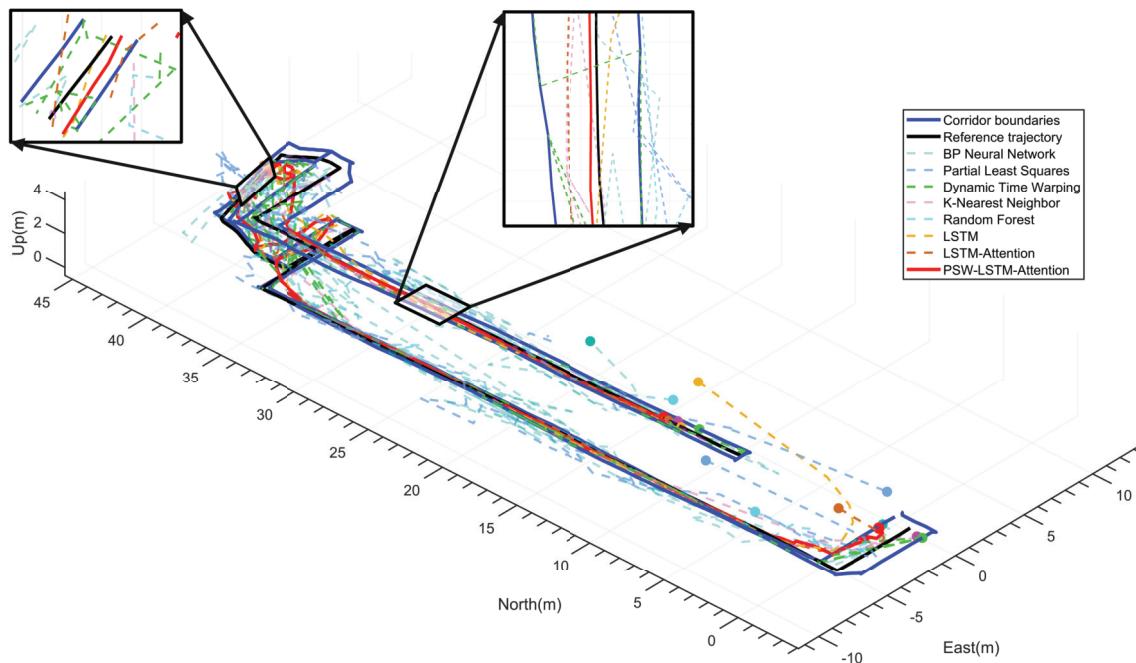


Fig. 9. Trajectory plot and zoomed-in trajectory plot of real-time data collected from the East Annex of the Library under different algorithms.

first phase to evaluate the quality of the fingerprint database construction and to compare the performance of different algorithms.

In the East Annex of the Library, the average corridor width is 2.5 m, with the long side measuring approximately 40.7 m and the short side approximately 6.3 m. The horizontal length of the staircase is around 10 m, and the floor height is 3.6 m. The total walking distance for training data (along the walls) is approximately 255 m, while for the test data (along the middle of the corridor), the distance is about 108.1 m. A new set of test data was collected two months later along the middle of the corridor, with a total walking distance of approximately 100.8 m.

Throughout the entire process, a total of 326 data points were collected for the training set. Using the DS-IDW interpolation method, the data points were supplemented to 640. During this process, Wi-Fi signals from 174 different

MAC addresses were collected. This resulted in a  $640 \times 178$  multisignal strength matrix (including Wi-Fi, magnetic field, and air pressure) as the input for the training set and a  $640 \times 3$  3-D position coordinate matrix as the output for the training set. The first test set collected 172 data points, forming a  $172 \times 178$  multisignal strength matrix and outputting the corresponding  $172 \times 3$  position coordinate matrix. The second test set, collected two months later, comprised 132 data points, forming a  $132 \times 178$  multisignal strength matrix and also outputting the corresponding 3-D position coordinates.

### *C. Experiment in the New Main Building*

Although the previous experiment provided an initial validation of the positioning performance in a 3-D indoor scene, the overall environment was relatively simple and may not fully represent the actual indoor positioning performance.

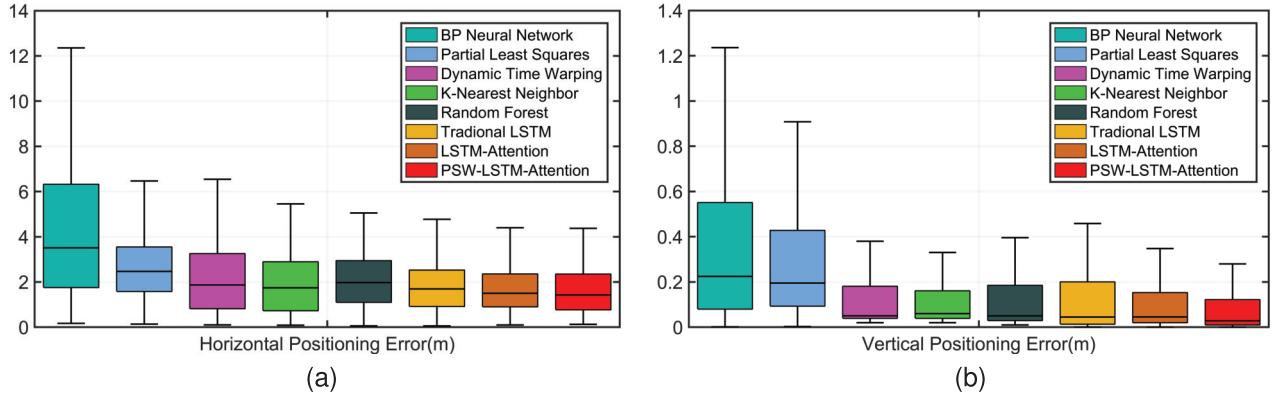


Fig. 10. Error bar for the first collection in the East Annex of the Library by different methods. (a) Horizontal positioning error. (b) Vertical positioning error.

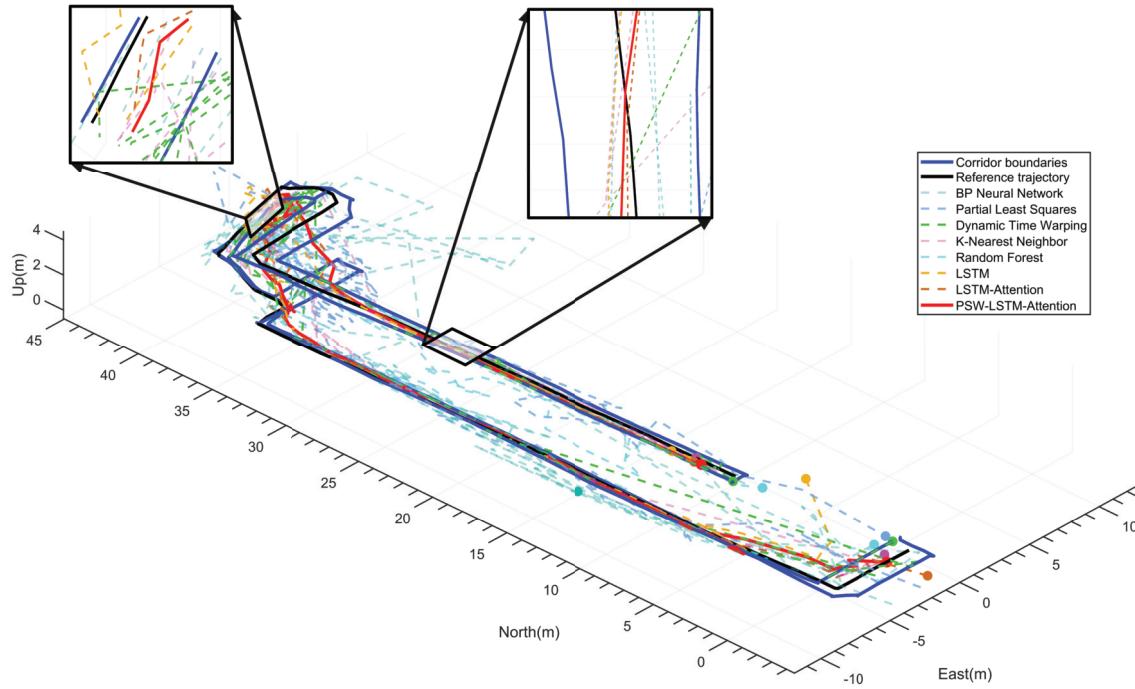


Fig. 11. Trajectory plot and zoomed-in trajectory plot of data collected two months later from the East Annex of the Library under different algorithms.

Therefore, the experiment was repeated on the seventh, eighth, and ninth floors of Building D in the New Main Building. This new setting includes multiple floors and various staircases and elevator scenarios, allowing for comprehensive testing in a more complex and diverse environment. Fig. 8 presents the specific data acquisition situation. Similarly, the cyan trajectory represents the path formed by connecting the reference points in the training set, while the yellow trajectory represents the path formed by connecting the points in the testing set.

As for the New Main Building, the average corridor width is also 2.5 m, with the long side measuring approximately 28.1 m and the short side 13 m. The horizontal length of the staircase between two floors is about 12 m, with a floor height of 4.2 m. The total walking distance for training data (along the walls) is 315.4 m, and for the test data (along the middle of the corridor), the distance is approximately 130.7 m.

For the training dataset, the data is first collected along the inner and outer walls of the corridor while moving and synchronizing data collection. Since the structure of each floor is essentially the same, the same method is applied on each floor. For the elevator section, a greater number of data points are collected while the elevator is stationary on each floor. During the operation of the elevator, data will be continuously collected. A total of 407 data points were collected, identifying 211 different Wi-Fi MAC addresses, and the corresponding air pressure and magnetic field strength were recorded synchronously. Subsequently, the DS-IDW method is used for data interpolation to supplement the dataset, resulting in 693 new data points. Thus, the input of the training set is an  $1100 \times 215$  matrix, and the output is an  $1100 \times 3$  matrix. For the test dataset, the data is collected while walking along the middle parts of the corridor and staircase trajectories. The collection starts from the midpoint of the seventh floor, proceeds to the left staircase, ascends to the eighth floor, walks along the

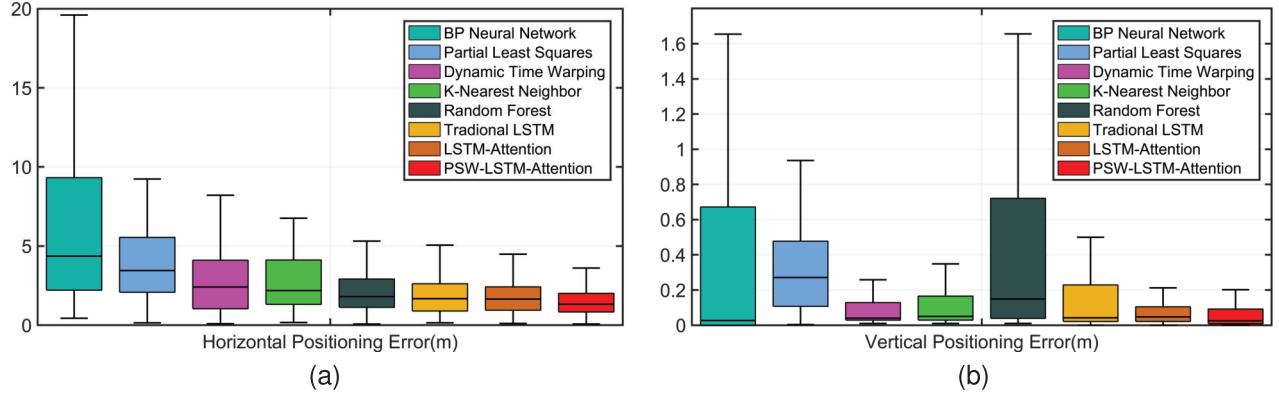


Fig. 12. Error bar for the recollection in the East Annex of the Library after two months by different methods. (a) Horizontal positioning error. (b) Vertical positioning error.

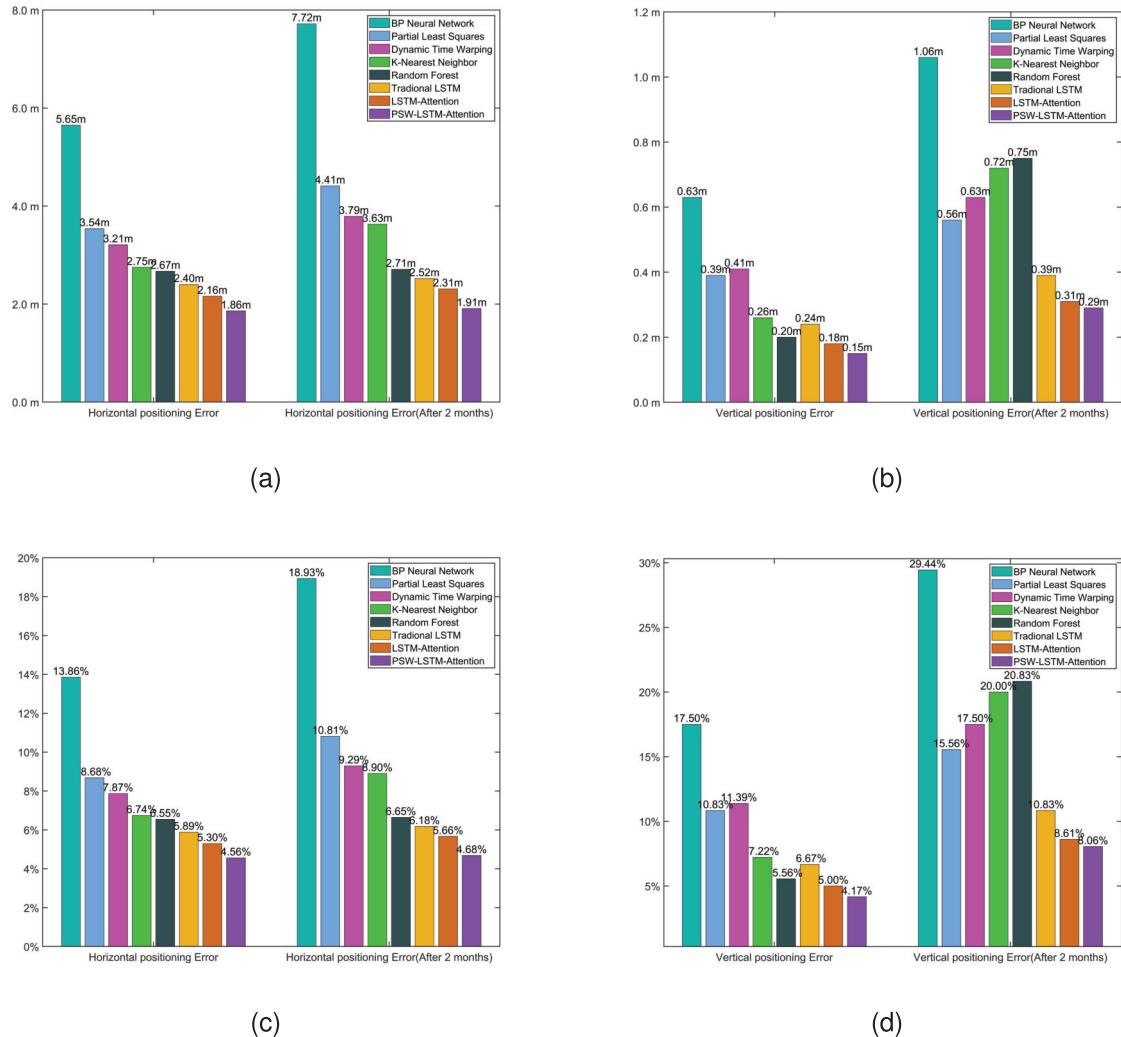


Fig. 13. (a) and (b) Absolute errors (i.e., RMSE) in the horizontal and vertical directions, respectively, in the Library East Annex experiment. (c) and (d) Relative errors (the absolute error divided by the actual indoor dimensions) in the horizontal and vertical directions.

corridor to the right staircase, ascends to the ninth floor, walks along the corridor to the middle elevator, and finally takes the elevator back to the starting point. A total of 119 data points are collected, forming a  $119 \times 215$  matrix.

This method covers almost all walkable areas on the floor plan and includes all common indoor scenarios, allowing for comprehensive verification of indoor positioning effectiveness. Additionally, the test path, which involves walking

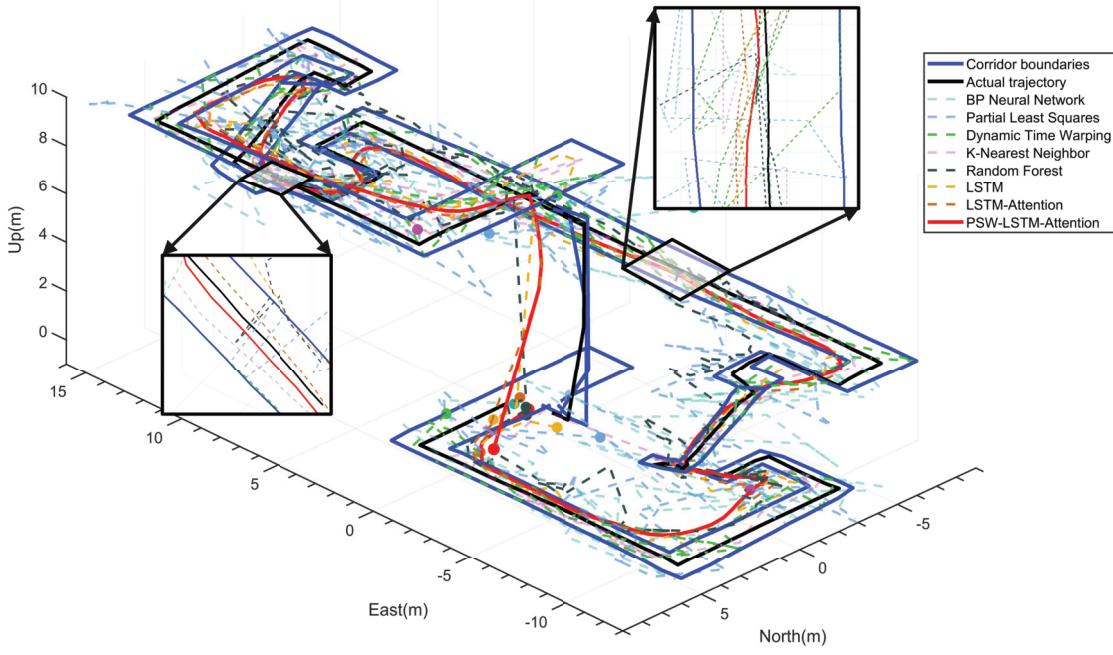


Fig. 14. Trajectory plot and zoomed-in trajectory plot of data collected from the New Main Building under different algorithms.

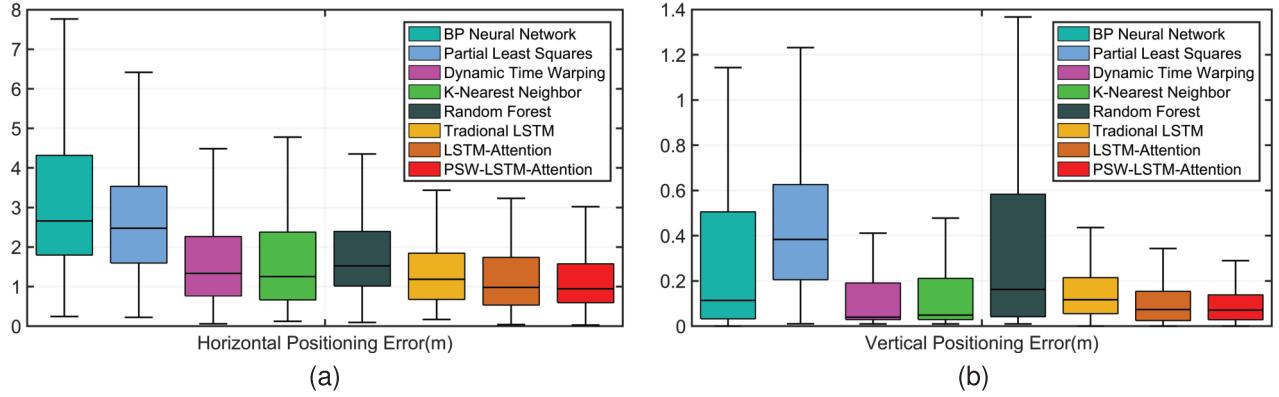


Fig. 15. Error bar for the collection in the New Main Building by different methods. (a) Horizontal positioning error. (b) Vertical positioning error.

and collecting data along the middle of the corridor, closely matches the typical walking paths of most pedestrians, thereby accurately representing the positioning performance during actual usage. Finally, the predicted position coordinates will be compared with the collected position coordinates to calculate the positioning error and evaluate the positioning performance. It is also important to note that in the aforementioned experiments, the algorithmic model maintained an identical architecture and parameter configuration across different times and scenarios, thereby demonstrating its transferability across diverse application contexts.

To promote further research and facilitate advancements in indoor positioning systems, the dataset will be made publicly available. This open-source dataset includes both the training and testing data, along with detailed annotations of the Wi-Fi signal strengths, magnetic field measurements, air pressure readings, and corresponding position coordinates. By sharing this data, we aim to provide a valuable resource for the research community to validate, compare, and improve upon

existing methodologies, thereby contributing to the ongoing development of robust and accurate indoor positioning solutions.

#### IV. EXPERIMENTAL RESULTS

To more intuitively demonstrate the advantages of PSW-LSTM with self-attention in positioning performance, this article will compare the performance of different machine learning models under the aforementioned experimental conditions. Specifically, the comparison includes the back-propagation neural network (BPNN) [29], PLS [30], random forest (RF) [33], DTW [31], KNNs [32], traditional LSTM [34], LSTM-Self-Attention, and the proposed improved PSW-LSTM with self-attention.

To promote further research and facilitate advancements in indoor positioning systems, the dataset will be made publicly available. This open-source dataset includes both the training and testing data, along with detailed annotations of the Wi-

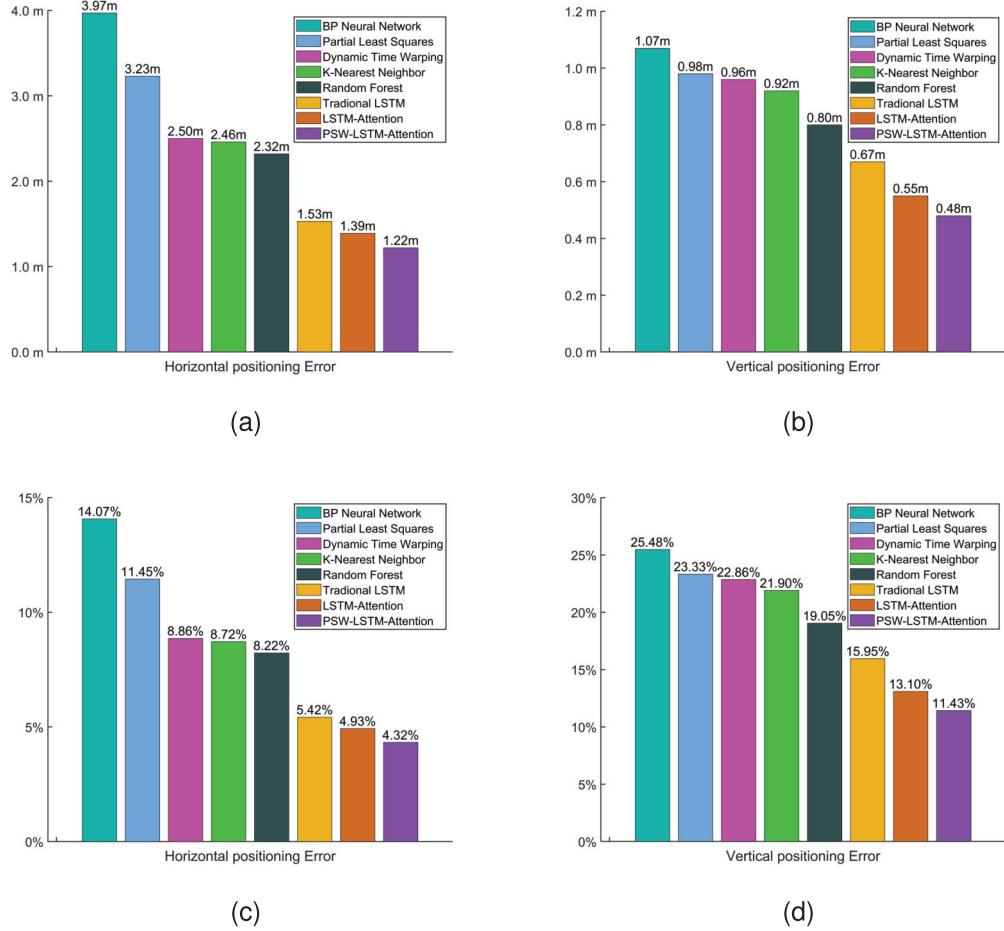


Fig. 16. (a) and (b) Absolute errors in the horizontal and vertical directions, respectively, in the New Main Building experiment. (c) and (d) Relative errors in the horizontal and vertical directions.

Fi signal strengths, magnetic field measurements, air pressure readings, and corresponding position coordinates. By sharing this data, we aim to provide a valuable resource for the research community to validate, compare, and improve upon existing methodologies, thereby contributing to the ongoing development of robust and accurate indoor positioning solutions.

#### A. Experiment in the East Annex of the Library

In this experiment, we specifically compared the performance of three different models on the test sets collected two months apart. Figs. 9 and 10 present the test results and error bar graphs in the horizontal and vertical directions, respectively, based on the real-time dataset collected from the East Annex of the library. Figs. 11 and 12 show the test results and error bar charts on the dataset collected again two months later. These figures specifically compare the positioning performance of the proposed method against BPNN, PLS, DTW, KNN, RF, LSTM, and LSTM-Self-Attention.

In these figures, the blue lines represent the trajectory formed by collecting data while walking along the sides of the corridor. The red lines represent the actual trajectory collected while walking in the middle of the corridor for the test set. The other trajectory represents the predicted trajectory obtained in

the test set. The dot dots indicate the starting and ending points of the matching trajectory.

Based on the results from the experiments, we can observe the following detailed results in Fig. 13.

Compared to BPNN, PLS, DTW, KNN, RF, traditional LSTM, and LSTM-Self-Attention, the overall performance of PSW-LSTM with self-attention on the real-time collected test set shows an improvement of 67.1%, 47.5%, 42.1%, 32.4%, 30.3%, 22.5%, and 13.9% in the horizontal and 76.2%, 61.5%, 63.4%, 42.3%, 25.0%, 37.5%, and 16.7% in the vertical directions, respectively, achieving accuracies of 1.86 and 0.15 m. On the test set collected two months later, the horizontal improved by 75.3%, 56.7%, 49.6%, 47.4%, 29.5%, 24.2%, and 17.3% and vertical accuracy improved by 72.6%, 48.2%, 54.0%, 59.7%, 61.3%, 25.6%, and 6.5% respectively, achieving accuracies of 1.91 and 0.29 m. Regardless of the timing, the method proposed in this article achieves the optimal positioning accuracy. In addition, compared to the real-time collected test set, the positioning performance in the test set collected two months later has decreased, especially in the vertical positioning accuracy using random forest, with an overall performance drop of over 70%. Moreover, based on the results of the relative error, the positioning errors for data collected at different times are all around 5%.

The method proposed in this article effectively mitigates this performance degradation, ensuring the long-term effectiveness of the signal fingerprint database. The overall positioning performance remains optimal, maintaining the functionality of high-precision indoor positioning.

### B. Experiment in the New Main Building

The overall experimental environment is more complex, encompassing multiple corridors, two different staircases, elevators, and multiple floors, which presents greater challenges for achieving precise indoor positioning. Fig. 14 shows the positioning results using BPNN, PLS, DTW, KNN, RF, LSTM, LSTM-Self-Attention, and PSW-LSTM with self-attention, respectively. Fig. 15 presents the error bar graphs for the positioning errors. Based on the results, it can be observed in Fig. 16.

The positioning accuracy using PSW-LSTM with self-attention reaches 1.22 m in the horizontal direction and 0.48 m in the vertical direction. This represents improvements of 69.3%, 62.2%, 51.2%, 50.4%, 47.4%, 20.3%, and 12.2% in the horizontal direction and 55.1%, 51.0%, 50.0%, 47.8%, 40.0%, 28.4%, and 12.7% in the vertical direction, compared to BPNN, PLS, DTW, KNN, RF, traditional LSTM, self-attention-enhanced LSTM, respectively.

Based on the results of the two sets of experiments, it can be seen that the proposed PSW-LSTM with self-attention method performs well in diverse and complex indoor environments. On the one hand, the introduction of the self-attention mechanism further enhances the performance of the LSTM model by flexibly selecting the most relevant information when dealing with diverse and complex inputs. On the other hand, the application of the PSW method can further reduce the issue of initial prediction offset. By organically unifying these two approaches, the overall predictive performance of the model can be further improved.

The error bar charts provide a clear visual representation of the model's performance across various test conditions. Notably, the proposed method demonstrates not only higher accuracy but also reduced variance compared to baseline methods. The error bars for the PSW-LSTM with the self-attention model are consistently smaller, indicating a more stable and reliable performance across different scenarios. This is particularly evident in complex environments, where traditional methods exhibit larger error margins and greater variability. Whether data is collected in real time for the test set or recollected after a period of time for testing, the method consistently achieves good positioning accuracy with lower error variance, reinforcing its robustness.

Additionally, this method demonstrates strong performance in large-scale 3-D spaces that include corridors, staircases, and elevators, achieving improved accuracy. The reduced variance further highlights the model's ability to maintain consistent performance in dynamic and challenging environments, such as stairways and corridors, where traditional methods struggle to maintain reliable accuracy.

## V. CONCLUSION AND FUTURE WORK

In this article, we proposed a novel framework that enhances indoor positioning efficiency by leveraging the collaboration between foot-mounted positioning terminals and smartphones. This feature fusion framework enables efficient sparse signal data collection, facilitating the rapid construction of a robust signal fingerprint database. By combining the strengths of both devices, the framework improves database quality and coverage, addressing limitations in traditional methods. We introduce a self-attention-enhanced PSW-LSTM model, significantly improving the accuracy of the positioning system. The self-attention mechanism enhances the model by accelerating convergence, enabling faster initialization of matching trajectories and improving the adaptability of the system to complex environments. These improvements contribute to higher robustness, allowing the model to maintain precise positioning even in dynamic and variable indoor environments where traditional methods struggle. This makes the framework highly suitable for applications requiring high accuracy and stability under challenging conditions.

Extensive experiments using real-world datasets across various complex environments showed that our method consistently achieved horizontal errors below 2 m and vertical errors under 0.5 m. The model demonstrated strong stability, maintaining high performance even with datasets collected months apart. To further support research in this field, the signal fingerprint database and testing data have been made publicly available, with plans for continuous updates, offering a valuable resource for ongoing experimentation and validation of positioning technologies.

This work provides a scalable, accurate solution for indoor positioning, reducing data collection efforts while improving performance. Future research will explore real-time application in dynamic environments and integration with additional sensors for broader use in autonomous systems, smart buildings, and augmented reality.

For future work, we plan to focus on two main areas of improvement.

- 1) Addressing the initial calibration issue: we aim to shorten the calibration time as much as possible, ensuring a faster and smoother user experience. By reducing the initial setup period, users will be able to use the system seamlessly without significant delays, making the application more practical for real-world scenarios.
- 2) Verifying algorithm robustness in diverse environments: we intend to test and validate the robustness of our algorithm in a wider range of environments, particularly in spaces that experience large dynamic changes, such as shopping malls or airports. These types of environments present unique challenges, and it is crucial to assess how well our algorithm can handle these dynamic conditions while maintaining high accuracy in localization.

## ACKNOWLEDGMENT

The authors would like to thank the editors and reviewers for their valuable comments, which significantly improved the quality of the manuscript.

## REFERENCES

- [1] J. Wang et al., "Seamless indoor-outdoor foot-mounted inertial pedestrian positioning system enhanced by smartphone PPP/3-D map/barometer," *IEEE Internet Things J.*, vol. 11, no. 7, pp. 13051–13069, Apr. 2024.
- [2] X. Li, "GNSS repeater based differential indoor positioning with multi-epoch measurements," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 1, pp. 803–813, Jan. 2023.
- [3] Y. You and C. Wu, "Hybrid indoor positioning system for pedestrians with swinging arms based on smartphone IMU and RSSI of BLE," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 9510615.
- [4] Y. Wang, J. Kuang, Y. Li, and X. Niu, "Magnetic field-enhanced learning-based inertial odometry for indoor pedestrian," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 2512613.
- [5] W. Zhang et al., "Cooperative positioning method of dual foot-mounted inertial pedestrian dead reckoning systems," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 8502114.
- [6] X. Liu, W. Zhang, J. Chen, D. Wei, and G. Yang, "NR-VINS: A real-time NIR-RGB visual-inertial odometry in challenging illumination environments," *IEEE Sensors J.*, vol. 25, no. 5, pp. 9070–9080, Mar. 2025.
- [7] D. Yan, T. Li, and C. Shi, "Enhanced online calibration and initialization of visual-inertial SLAM system leveraging the structure information," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 5029715.
- [8] S. Wen, S. Tao, X. Liu, A. Babiarz, and F. R. Yu, "CD-SLAM: A real-time stereo visual-inertial SLAM for complex dynamic environments with semantic and geometric information," *IEEE Trans. Instrum. Meas.*, vol. 73, 2024, Art. no. 2517808.
- [9] A. Perttula, H. Leppäkoski, M. Kirkko-Jaakkola, P. Davidson, J. Collin, and J. Takala, "Distributed indoor positioning system with inertial measurements and map matching," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 11, pp. 2682–2695, Nov. 2014.
- [10] C.-H. Huang, L.-H. Lee, C. C. Ho, L.-L. Wu, and Z.-H. Lai, "Real-time RFID indoor positioning system based on Kalman-filter drift removal and heron-bilateration location estimation," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 728–739, Mar. 2015.
- [11] T. Liu, X. Niu, J. Kuang, S. Cao, L. Zhang, and X. Chen, "Doppler shift mitigation in acoustic positioning based on pedestrian dead reckoning for smartphone," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.
- [12] B. Yang, J. Li, and H. Zhang, "Resilient indoor localization system based on UWB and visual-inertial sensors for complex environments," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 8504014.
- [13] Q. Tian, K. I.-K. Wang, and Z. Salcic, "A resetting approach for INS and UWB sensor fusion using particle filter for pedestrian tracking," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 8, pp. 5914–5921, Aug. 2020.
- [14] P. S. Farahsari, A. Farahzadi, J. Rezaizadeh, and A. Bagheri, "A survey on indoor positioning systems for IoT-based applications," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7680–7699, May 2022.
- [15] N. Xiaojie, W. Yan, and K. Jian, "A pedestrian POS for indoor mobile mapping system based on foot-mounted visual-inertial sensors," *Measurement*, vol. 199, Aug. 2022, Art. no. 111559.
- [16] J. Kuang, T. Li, Q. Chen, B. Zhou, and X. Niu, "Consumer-grade inertial measurement units enhanced indoor magnetic field matching positioning scheme," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, 2023.
- [17] Y. Wang, Y. Wang, Q. Liu, and Y. Zhang, "Dynamic WiFi indoor positioning based on the multi-scale metric learning," *Comput. Commun.*, vol. 213, pp. 49–60, Jan. 2024.
- [18] D. B. Ninh, J. He, V. T. Trung, and D. P. Huy, "An effective random statistical method for indoor positioning system using WiFi fingerprinting," *Future Gener. Comput. Syst.*, vol. 109, pp. 238–248, Aug. 2020.
- [19] X. Kong, C. Wu, Y. You, Z. Lv, and Z. Zhao, "Hybrid indoor positioning method of BLE and monocular VINS-based smartphone," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–13, 2023.
- [20] J. Chen et al., "A data-driven inertial navigation/Bluetooth fusion algorithm for indoor localization," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5288–5301, Mar. 2022.
- [21] M. Ji, X. Xu, and Y. Guo, "An adaptive heading correction algorithm for suppressing magnetic interference in inertial navigation system," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.
- [22] G. De Angelis et al., "An indoor AC magnetic positioning system," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 5, pp. 1267–1275, May 2015.
- [23] F. Zhao, H. Luo, X. Zhao, Z. Pang, and H. Park, "HYFI: Hybrid floor identification based on wireless fingerprinting and barometric pressure," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 330–341, Feb. 2017.
- [24] X. Tong, Y. Wan, Q. Li, X. Tian, and X. Wang, "CSI fingerprinting localization with low human efforts," *IEEE/ACM Trans. Netw.*, vol. 29, no. 1, pp. 372–385, Feb. 2021.
- [25] J. Lu et al., "ONavi: Data-driven based multi-sensor fusion positioning system in indoor environments," in *Proc. IEEE 12th Int. Conf. Indoor Positioning Indoor Navig. (IPIN)*, Beijing, China, Sep. 2022, pp. 1–8.
- [26] J. Tan, H. Wu, K.-H. Chow, and S.-H. G. Chan, "Implicit multimodal crowdsourcing for joint RF and geomagnetic fingerprinting," *IEEE Trans. Mobile Comput.*, vol. 22, no. 2, pp. 935–950, Feb. 2023.
- [27] F. Gu et al., "Accurate and efficient floor localization with scalable spiking graph neural networks," *Satell. Navig.*, vol. 5, no. 1, pp. 6–21, Feb. 2024.
- [28] N. Zhu, H. Zhao, W. Feng, and Z. Wang, "A novel particle filter approach for indoor positioning by fusing WiFi and inertial sensors," *Chin. J. Aeronaut.*, vol. 28, no. 6, pp. 1725–1734, Dec. 2015.
- [29] Y. Zhao, T. Wang, Q. Miao, J. Yan, and R. Wang, "Research on indoor and outdoor positioning switching algorithm based on improved PSO-BP," *Meas. Sci. Technol.*, vol. 35, no. 8, Aug. 2024, Art. no. 086313.
- [30] C. Chen, Y. Wang, Y. Zhang, and Y. Zhai, "Indoor positioning algorithm based on nonlinear PLS integrated with RVM," *IEEE Sensors J.*, vol. 18, no. 2, pp. 660–668, Jan. 2018.
- [31] K. Qiu, H. Huang, W. Li, and D. Luo, "Indoor geomagnetic positioning based on a joint algorithm of particle filter and dynamic time warp," in *Proc. Ubiquitous Positioning, Indoor Navig. Location-Based Services (UPINLBS)*, Wuhan, China, Mar. 2018, pp. 1–7.
- [32] H. Zhang, Z. Wang, W. Xia, Y. Ni, and H. Zhao, "Weighted adaptive KNN algorithm with historical information fusion for fingerprint positioning," *IEEE Wireless Commun. Lett.*, vol. 11, no. 5, pp. 1002–1006, May 2022.
- [33] W. Ji, K. Zhao, Z. Zheng, C. Yu, and S. Huang, "Multivariable fingerprints with random forest variable selection for indoor positioning system," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5398–5406, Mar. 2022.
- [34] M. Zhang, J. Jia, J. Chen, L. Yang, L. Guo, and X. Wang, "Real-time indoor localization using smartphone magnetic with LSTM networks," *Neural Comput. Appl.*, vol. 33, no. 16, pp. 10093–10110, Aug. 2021.
- [35] Q. Wang et al., "Recent advances in pedestrian inertial navigation based on smartphone: A review," *IEEE Sensors J.*, vol. 22, no. 23, pp. 22319–22343, Dec. 2022.
- [36] M. T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp, and K. Reddy, "Recurrent neural networks for accurate RSSI indoor localization," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10639–10651, Dec. 2019.
- [37] G. Brauwelaars and F. Frasincar, "A general survey on attention mechanisms in deep learning," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 3279–3298, Apr. 2023.
- [38] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5929–5955, May 2020.
- [39] Q. Guo et al., "Motion pattern recognition for indoor pedestrian altitude estimation based on inertial sensor," *IEEE Sensors J.*, vol. 24, no. 6, pp. 8197–8209, Mar. 2024.
- [40] M. Xia, J. Wang, C. Shi, and W. Wen, "Indoor altitude estimation assisted by inertial compensation and online floor modeling," *Meas. Sci. Technol.*, vol. 35, no. 12, Sep. 2024, Art. no. 126302.
- [41] J. Wang et al., "Multi-frequency smartphone positioning performance evaluation: Insights into A-GNSS PPP-B2b services and beyond," *Satell. Navigat.*, vol. 5, no. 1, pp. 1–26, Aug. 2024.
- [42] X. Zhao, Y. Meng, F. Qi, L. Wang, and X. Zhu, "A vertical channel-enhanced fusion method based on RINS and barometric altimeter for UAVs in GNSS denial environments," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023.
- [43] L. Cong, J. Tian, and H. Qin, "A practical hybrid height estimation algorithm based on FM-aided motion mode recognition," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–13, 2023.
- [44] J. Guo, P. Li, H. Wu, and S. Hu, "An analytical model of the electric field distribution and breakdown voltage for stepped compound buried layer SOI LDMOS," in *Proc. Int. Conf. Electr. Comput. Energy Technol. (ICECET)*, Prague, Czech Republic, Jul. 2022, pp. 1–6.
- [45] T. Hidayat and I. A. Astuti, "Interactive augmented reality for the depth of an object using the model-based occlusion method," in *Proc. 3rd Int. Conf. Comput. Informat. Eng. (IC2IE)*, Yogyakarta, Indonesia, Sep. 2020, pp. 382–387.
- [46] M. Liu, J. Fan, Y. Zheng, S. Li, and L. Jin, "A simultaneous learning and control scheme for redundant manipulators with physical constraints on decision variable and its derivative," *IEEE Trans. Ind. Electron.*, vol. 69, no. 10, pp. 10301–10310, Oct. 2022.



**Zhanpeng Zhang** is currently pursuing the M.S. degree with the School of Electronics and Information Engineering, Beihang University, Beijing, China.

His current research interests include multisource fusion navigation, pedestrian localization, and indoor positioning, which include GNSS, inertial, and fingerprint navigation.



**Jiale Wang** (Member, IEEE) received the bachelor's degree in communication engineering from Beijing Jiaotong University, Beijing, China, in 2019, and the Ph.D. degree in communication and information systems from Beihang University, Beijing, in 2024.

He is currently an Assistant Researcher at Beihang University. His current research focuses on high-precision GNSS positioning, multisensor fusion, and the Internet of Things.



**Chuang Shi** received the Ph.D. degree from Wuhan University, Wuhan, China, in 1998.

He currently serves as a Professor with the School of Electronic and Information Engineering, Beihang University, Beijing, China. Additionally, he also holds the position of the Director of the Key Laboratory of Satellite Navigation and Mobile Communication Fusion Technology, a Key Laboratory under the Ministry of Industry and Information Technology. He has been distinguished with the National Science Fund for Distinguished Young

Scholars and holds the esteemed title of Distinguished Professor within the Chang Jiang Scholars Program. He has led the development of China's autonomous software systems for precision orbit determination and positioning of navigation satellites, as well as the wide-area real-time precise positioning system and the critical technology underpinning the BDS ground-based augmentation system, contributing to several national key Research and Development initiatives. His research contributions have been successfully integrated into major projects, such as the "BDS Navigation Major Project," the "XiHe System," and the "National BDS Ground-Based Augmentation System." His prolific academic output includes authoring or coauthoring more than 100 academic papers, securing 20 patents, and holding copyrights for five pieces of software. His research interests include network adjustment, precise orbit determination of GNSS satellites and LEOs, as well as real-time PPP.

Prof. Shi's exemplary work has been recognized with two first-place National Science and Technology Progress Prizes and more than ten awards at the provincial or ministry level.



**Weisong Wen** (Member, IEEE) received the Ph.D. degree in mechanical engineering from The Hong Kong Polytechnic University, Hong Kong, in 2020.

He was a Visiting Student Researcher at the University of California, Berkeley (UCB), Berkeley, CA, USA, in 2018. He is currently a Research Assistant Professor with the Department of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University. His research interests include multisensor integrated localization for autonomous vehicles, SLAM, and GNSS positioning in urban canyons.



**Deyou Zhang** (Member, IEEE) received the B.S. and M.S. degrees from Harbin Institute of Technology, Harbin, China, in 2012 and 2014, respectively, and the Ph.D. degree from the School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW, Australia, in 2020.

He currently serves as an Associate Professor with the School of Electronic and Information Engineering, Beihang University, Beijing, China. His research interests include millimeter-wave communications, reconfigurable intelligent surfaces, and wireless edge



**Ming Xia** received the Ph.D. degree in signal and information processing from Beihang University, Beijing, China, in 2020.

He is currently a Post-Doctoral Fellow with the School of Electronic and Information Engineering, Beihang University. His research interests include motion recognition, pedestrian inertial positioning, wearable sensor-based positioning, and their applications in location-based service applications, especially in the firefighting field.

intelligence.