IPIN 2025 Competition – Competition Technology Report

1. Team Information

- Team Name: IPIN_CUMT
- Affiliation: School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China
- Country: China
- Track: Track 3 “Smartphone (Offsite-Online)”
- Team Members:
  Zengke Li; Meng Sun; Yiruo Li; Qi wu; Pengkai He; Shiyi Jing and Yiting Wu

2. Technological Route

Fig. 1 presents the general framework of our proposed positioning system for the IPIN 2025 competition.

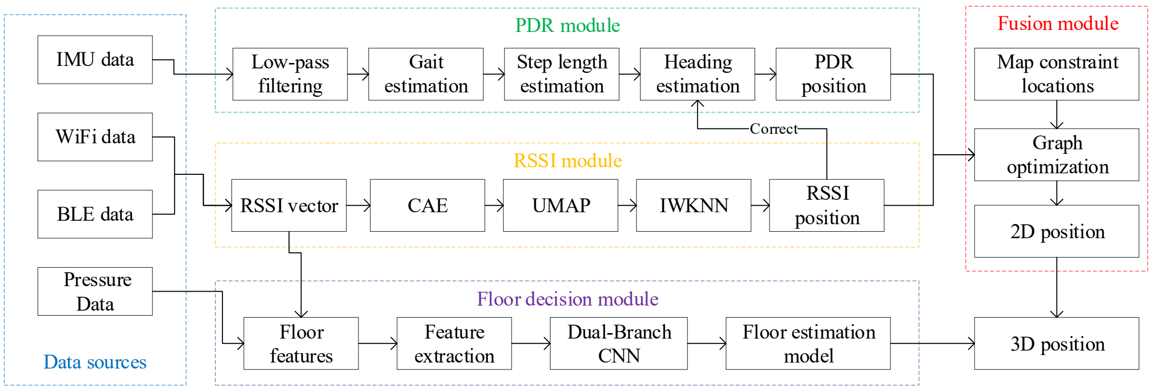

Fig.1. The framework of the proposed positioning system.

The proposed positioning system contains four main sections: data sources; PDR module; RSSI module; floor decision and fusion module

1) Data sources; this section describes the positioning data sources utilized by our proposed localization system, including IMU, WiFi, BLE, and barometric pressure sensors.

- IMU data includes accelerometer, gyroscope and magnetometer signals as well as attitude data solved by the smartphone.
- WiFi data: RSSI values from surrounding access points; serves as the basis for fingerprint localization.
- Bluetooth signal strength values, used similarly to WiFi for proximity-based positioning.
- The pressure data, obtained from a barometric sensor, is used to support floor-level estimation.

2) RSSI module; this module performs localization using WiFi/BLE signal fingerprints with a dimensionality reduction and matching pipeline:

- RSSI vector: This is constructed by concatenating the RSSI values from nearby Wi-Fi access points and Bluetooth beacons. It captures the multi-source wireless signal fingerprint used for position estimation.
- CAE (Convolutional Autoencoder): Its main purpose is to use CAE's hierarchical feature learning to derive a preliminary low-dimensional embedding, that is, to perform preliminary feature extraction on the RSSI space. In the process of training the CAE model, we added a triplet loss to

force the model to learn discriminative features.

- UMAP (Uniform Manifold Approximation and Projection): A nonlinear manifold learning method that preserves global data distribution and local structure by modeling nonlinear topology. Benefiting from the compact feature representation initially extracted by the encoder, UMAP operates on a cleaner and more informative input space, further reducing the feature dimensionality and obtaining discriminative low-dimensional embeddings.
- IWKNN (Improved Weighted K-Nearest Neighbors): An improved WKNN algorithm is designed, combining feature weighting and dynamic k value selection to achieve robust location estimation. Namely, we incorporate the dimensionality variance as an adaptive weighting factor into the Euclidean distance calculation and apply a dynamic threshold-based filtering mechanism to exclude geometrically inconsistent candidates. Finally, the low-dimensional feature representation output by UMAP is matched with the candidate offline reference fingerprint to achieve robust location estimation.

3) PDR (Pedestrian Dead Reckoning) module; the PDR module estimates the user's relative displacement using IMU signals through the following pipeline:

- Low-pass filtering: Removes high-frequency noise from raw IMU data.
- Gait estimation: Detects steps based on acceleration patterns (e.g., zero-crossing or peak detection).
- Step length estimation: Uses empirical or learning-based models to calculate step size per stride.
- Heading estimation: Pedestrian heading estimation is achieved by integrating measurements from the accelerometer, gyroscope, and magnetometer. In addition, considering that the smartphone's holding posture can significantly affect the heading estimation, we incorporate valid Wi-Fi positioning results to constrain the heading estimation. Specifically, the heading is corrected by fitting the slope of consecutively forward-moving Wi-Fi positioning results.

4) Floor decision module; this module is responsible for identifying the correct floor level of the user, which is critical in multi-floor buildings:

- Floor features: Constructed by integrating pressure readings and WiFi/BLE RSSI vector.
- Feature extraction: Preprocessing and transformation into a compact vector representation.
- Dual-Branch CNN: A deep neural network with two branches for fusing floor features (e.g., pressure + RSSI).
- Floor Estimation Model: Leveraging the real-time RSSI vector and barometric pressure data, the model effectively estimates the current floor level where the user is located.

5) Fusion module; the fusion module combines the outputs from the PDR and RSSI modules, enhanced by map constraints and graph-based optimization:

- Map constraint locations: If reliable map information is available, candidate locations such as walkable areas and doorways can be extracted from the building map. Then, RSSI measurements are used to determine whether the pedestrian has reached these candidate locations. If so, the candidate locations are used as constraints to refine the positioning result.
- Graph optimization: Builds a pose graph where nodes represent estimated positions and edges represent relative displacements or map constraints.
- 2D position: The final refined horizontal position estimate.
- 3D position: Combining 2D location with predicted floor level from the floor decision module to yield full 3D indoor position.

3. Construction of fingerprint database

For the RSSI module, we primarily implement localization using a fingerprint-based approach.

However, the training data provided by the IPIN competition does not include the precise coordinates corresponding to each RSSI measurement. In this section, we present our method for constructing the RSSI fingerprint database based on the available training data. The detailed procedure for building the offline RSSI fingerprint database is illustrated in Fig. 2.

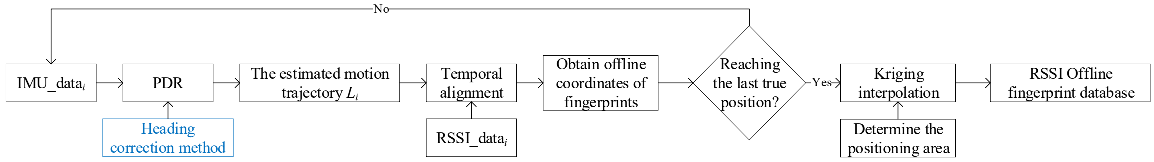

Fig.2. The process of constructing the offline RSSI fingerprint database. Here IMU_data$_i$ and RSSI_data$_i$ represent the IMU and RSSI data collected between the i th and (i+1)th ground truth positions, respectively.

The provided training trajectory data contains only a limited number of ground truth coordinates. To address this limitation, a PDR is employed to process the IMU data collected between consecutive ground truth locations (i and i+1), generating a dense and continuous trajectory. The corresponding RSSI measurements are temporally aligned with the PDR-estimated positions using synchronized timestamps, forming initial location-RSSI fingerprint pairs. To ensure sufficient spatial coverage of the environment, spatial interpolation techniques (e.g., Kriging) are applied to these sparse fingerprints, estimating RSSI values at unvisited locations. This process results in a dense, grid-based fingerprint database that maps physical coordinates to signal strength vectors across the entire target area.

It is worth noting that the smartphone’s heading does not always accurately reflect the pedestrian’s actual movement direction. Therefore, we propose a heading correction method to constrain the estimated heading, making it more consistent with the pedestrian's true walking direction. Specifically, the constrained heading θ_h^i is obtained from the coordinates of the ith and (i+1)th true location. We then estimate the heading based primarily on accelerometer and magnetic data and use that constrained heading to constrain the estimated heading. Firstly, we rotate the acceleration data to the navigation coordinate system:

$$\begin{bmatrix} a_x^N \\ a_y^N \\ a_z^N \end{bmatrix} = \mathbf{R}_I^N \begin{bmatrix} \tilde{a}_x^I \\ \tilde{a}_y^I \\ \tilde{a}_z^I \end{bmatrix} \tag{1}$$

Here, $\mathbf{R}_I^N$ notes the rotation matrix form IMU frame to the navigation frame. $\left[\tilde{a}_x^I, \tilde{a}_y^I, \tilde{a}_z^I\right]$ is the acceleration data processed by low-pass filtering and removing the effects of gravity. Then, the estimated heading $\hat{\theta}_a$ using the acceleration data is calculated by:

$$\hat{\theta}_a = \arctan 2\left(a_y^N, a_x^N\right) \tag{2}$$

For the estimated heading by the magnetic data is calculated by:

$$\hat{\theta}_m = \arctan\left(\frac{m_x^I \cos\phi + m_z^I \sin\phi}{m_x^I \sin\rho \sin\phi + m_y^I \cos\theta - m_z^I \cos\rho \sin\phi}\right) \tag{3}$$

where, $\left[m_x^I, m_y^I, m_z^I\right]$ is the magnetic data processed by low-pass filtering, ϕ denotes the pitch and ρ denotes the roll.

Then, the final estimated heading of the pedestrian is calculated by:

$$\hat{\theta} = \alpha \cdot \hat{\theta}_a + \beta \cdot \hat{\theta}_m + (1 - \alpha - \beta) \cdot \theta_h^i \tag{4}$$

This method integrates acceleration-based direction estimation, magnetometer heading, and path-constrained heading derived from the pedestrian's trajectory to improve heading accuracy. By combining complementary sensor sources with trajectory-based constraints, it effectively mitigates the adverse effects of smartphone orientation errors, which often degrade heading estimation due to inaccurate pitch and roll. The trajectory-derived heading acts as a global constraint that stabilizes the estimated heading, particularly in straight-line walking scenarios, thereby reducing drift and compensating for momentary sensor noise or orientation fluctuations.

4. Summarize

This report presents a robust smartphone-based indoor positioning system developed by Team IPIN_CUMT for IPIN 2025 Track 3. The solution integrates Pedestrian Dead Reckoning (PDR) with multi-source wireless fingerprinting (WiFi/BLE) and barometric pressure sensing within a unified framework. Key innovations include: 1) A CAE-UMAP-IWKNN pipeline for discriminative RSSI fingerprint feature extraction, dimensionality reduction, and robust matching; 2) Enhanced PDR using acceleration/magnetometer data constrained by WiFi-derived path headings to mitigate device orientation errors; 3) A dual-branch CNN for floor-level decision by fusing pressure and RSSI features; 4) Sparse fingerprint database construction via PDR trajectory interpolation and Kriging spatial estimation; and 5) Final fusion through map-constrained graph optimization, generating accurate 3D positions by combining refined 2D coordinates with predicted floor levels. This multi-modal approach effectively addresses challenges in signal variability, device heterogeneity, and multi-floor environments.