EKF-Based Fusion of Wi-Fi/LiDAR/IMU for Indoor Localization and Navigation

Zeyi Li[1,2], Zhe Tang[3], Kyeong Soo Kim[1], Sihao Li[4], and Jeremy S. Smith[2]
[1]School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, P.R. China.
[2]Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, U.K.
[3]Zhejiang University of Technology, Hangzhou, 310014, P.R. China.
[4]School of Artificial Intelligence, Suzhou Vocational Institute of Industrial Technology, Suzhou, 215104, P.R. China.
Email: Zeyi.Li21@alumni.xjtlu.edu.cn, timtangprc@gmail.com, Kyeongsoo.Kim@xjtlu.edu.cn, 01177@siit.edu.cn, J.S.Smith@liverpool.ac.uk

***Abstract*—Conventional Wi-Fi received signal strength indicator (RSSI) fingerprinting cannot meet the growing demand for accurate indoor localization and navigation due to its lower accuracy, while solutions based on light detection and ranging (LiDAR) can provide better localization performance but is limited by their higher deployment cost and complexity. To address these issues, we propose a novel indoor localization and navigation framework integrating Wi-Fi RSSI fingerprinting, LiDAR-based simultaneous localization and mapping (SLAM), and inertial measurement unit (IMU) navigation based on an extended Kalman filter (EKF). Specifically, coarse localization by deep neural network (DNN)-based Wi-Fi RSSI fingerprinting is refined by IMU-based dynamic positioning using a Gmapping-based SLAM to generate an occupancy grid map and output high-frequency attitude estimates, which is followed by EKF prediction-update integrating sensor information while effectively suppressing Wi-Fi-induced noise and IMU drift errors. Multi-group real-world experiments conducted on the IR building at Xi'an Jiaotong-Liverpool University demonstrates that the proposed multi-sensor fusion framework suppresses the instability caused by individual approaches and thereby provides stable accuracy across all path configurations with mean two-dimensional (2D) errors ranging from 0.2449 m to 0.3781 m. In contrast, the mean 2D errors of Wi-Fi RSSI fingerprinting reach up to 1.3404 m in areas with severe signal interference, and those of LiDAR/IMU localization are between 0.6233 m and 2.8803 m due to cumulative drift.**

***Index Terms*—Indoor localization, Wi-Fi fingerprinting, simultaneous localization and mapping (SLAM), inertial measurement unit (IMU), multi-sensor data fusion, extended Kalman filter (EKF), deep neural networks (DNNs).**

I. Introduction

With the widespread deployment of indoor service robots, intelligent logistics systems, and mobile terminals in smart buildings, high-accuracy indoor localization technology has become a key enabler for space perception and autonomous navigation. However, traditional global positioning systems (GPS) are severely obstructed in indoor environments, where signal attenuation and multipath interference render them virtually useless [1]. Therefore, researchers have turned to indoor localization methods based on wireless signals and sensor fusion to accurately track moving objects in complex spatial environments.

Currently, mainstream indoor localization methods include Wi-Fi fingerprinting, ultra wide band (UWB) ranging, Bluetooth near-field localization, and inertial measurement unit (IMU) navigation [2]. Wi-Fi fingerprinting has been favored in many application scenarios because it does not require the installation of any new infrastructure or the modification of existing devices, which results in lower deployment cost. However, this method relies heavily on a pre-collected received signal strength indicator (RSSI) database, which is susceptible to environmental changes and device differences, resulting in poor stability and versatility [3], [4]. IMU can quickly estimate the orientation (attitude) and motion of a body and thereby is widely used in fast localization tasks in dynamic environments. Classic IMU localization methods include the double integration method based on physical models, the pedestrian dead reckoning (PDR) method based on path estimation, the Kalman filter algorithm, and fusion methods with other sensors [5]–[8]. However, the localization performance of these methods degrades quickly as the errors in the acceleration and angular velocity integration processes, especially the drift effect, accumulate [9]. Light detection and ranging (LiDAR) offers extremely high spatial ranging accuracy and is commonly used for map construction and pose estimation in simultaneous localization and mapping (SLAM) systems [10]. However, its high cost and environmental dependencies limit its scalability for large-scale deployment.

Due to the limitations of the single-sensor solutions mentioned above, multi-sensor fusion methods have garnered widespread attention, in recent years, because joint modeling and filter estimation of different sensor information can greatly improve the robustness and localization accuracy of a system [11]. Among them, the extended Kalman filter (EKF), an established method for processing non-linear state estimation, has been widely used in robot perception and navigation systems due to its advantages of strong real-time performance, clear structure, and simple engineering implementation [12], [13]. However, current research is mostly limited to two-source

Z. Tang and S. Li were with the School of Advanced Technology at Xi'an Jiaotong-Liverpool University and the Department of Electrical Engineering and Electronics at the University of Liverpool when this work was done.

* Corresponding author: Kyeongsoo.Kim@xjtlu.edu.cn

fusion schemes such as Wi-Fi/PDR or UWB/IMU [14].

We propose a multi-sensor fusion framework for accurate and dynamic indoor localization and navigation under complex indoor environments, where Wi-Fi RSSI fingerprinting and LiDAR SLAM enhanced by IMU, are integrated based on EKF. We carry out a comparative analysis of the performance of the proposed framework through comprehensive experiments in the IR building of Xi'an Jiaotong-Liverpool University (XJTLU), whose results demonstrate that the proposed framework consistently outperforms Wi-Fi and LiDAR/IMU-based solutions, achieving superior error control and trajectory smoothness performance.

The rest of the paper is organized as follows: Section II reviews related work. Section III introduces the proposed EKF-based indoor localization framework. Section IV presents experimental results and compares the performance of different localization methods across multiple environments. Section V presents the conclusions from our work.

II. Related Work

A. EKF in Indoor Localization

EKF is a recursive state estimation algorithm widely used in indoor localization systems [15]. It can fuse asynchronous, noisy, and multiple sensor observations in environments with non-linear dynamic characteristics and estimate user location information in real time within a probabilistic framework. Compared to traditional linear Kalman filters, the EKF linearizes non-linear systems using a first-order Taylor expansion, enabling it to handle non-linear motion and observation models. This feature makes it more adaptable and robust in indoor environments with complex user trajectories and uncertain signal propagation paths [16].

The performance of indoor localization systems based on Wi-Fi RSSIs and channel state information (CSI) can be significantly affected due to multipath effects, channel attenuation, and dynamic changes in the environment [17]. EKF can filter and correct noisy observations by modeling state transitions and observation processes. To further improve system performance, researchers have recently combined EKF with inertial and environmental perception modalities such as IMU and LiDAR to fuse absolute localization and relative motion information, thereby enhancing the continuity and stability of trajectory estimation [17].

In addition, some studies have introduced deep neural networks (DNNs) for preprocessing Wi-Fi fingerprint data (e.g., using autoencoders for feature denoising and reconstruction to reduce the impact of environmental factors on signal distribution). Subsequently, Wi-Fi features enhanced by DNNs are jointly input into the EKF framework along with IMU data, enabling the system to simultaneously perceive static scene features and dynamic motion information during state updates, thereby effectively mitigating initial heading errors and short-term drift issues [18].

B. Multi-Sensor Fusion in Indoor Localization

Conventional indoor localization methods based on a single signal source are susceptible to dynamic changes and channel interference in complex scenarios, reducing localization accuracy and system stability. For this reason, multi-sensor fusion has received widespread attention in recent years as a promising solution to improve the accuracy and robustness of indoor localization systems. This technology integrates heterogeneous sensor data from different modalities to achieve more accurate, continuous, and robust estimates of target locations, effectively addressing uncertainties in dynamic environments [19].

From the perspective of fusion architecture, multi-sensor fusion methods can be categorized into data-level, feature-level, and decision-level fusion [20]. Data-level fusion directly integrates raw observation data from sensors, exploiting the most complete information. However, it has high requirements for time synchronization and noise robustness. Feature-level fusion, on the other hand, integrates the extracted features of each sensor's data, balancing computational efficiency and localization accuracy. Decision-level fusion weights or filters the results of the inference by a subsystem dedicated to each sensor, which is suitable for distributed architectures and heterogeneous device environments. Examples of recent development in multi-sensor fusion in indoor localization include the complementary fusion of Wi-Fi and IMU to improve short-term stability in scenarios without GPS [21], cooperative localization of UWB and visual sensors (such as cameras or LiDAR) to achieve centimeter-level accuracy in industrial and robotics scenarios [22], and multi-modal fusion strategies based on end-to-end neural networks to improve robust perception capabilities in complex environments through non-linear modeling [23].

Although multi-sensor fusion can improve localization accuracy and stability, it still faces many challenges [20]. For example, issues such as data synchronization between multi-source sensors, coordinate system alignment, and scale consistency have not been fully resolved. Likewise, the limited computational power of mobile devices imposes serious challenges to the design of real-time, energy-efficient fusion algorithms. In addition, data heterogeneity, outlier handling, and the complexity of system integration further restrict the feasibility of such methods for large-scale deployment.

III. Methodology

We construct a high-accuracy and robust indoor localization system that integrates three types of sensor information: Wi-Fi RSSIs, LiDAR-based SLAM, and IMU three-axis acceleration and angular velocity. It achieves state estimation and real-time correction through EKF. The method covers four significant steps: Data collection, modeling and training, state estimation, and performance evaluation. The overview of the proposed system is shown in Fig. 1.

A. Data Processing and Integration Process

The proposed system is implemented on the autonomous guided vehicle (AGV) platform shown in Fig. 2. This platform

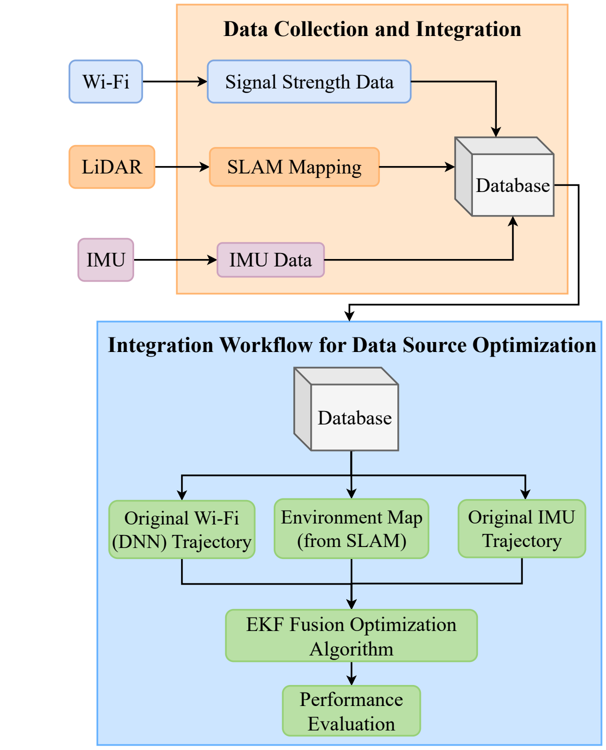


Fig. 1: Overall architecture and information processing flow of the multi-sensor indoor localization and navigation system.

integrates an NVIDIA Jetson computing module, a Wi-Fi receiver module, a two-dimensional (2D) LiDAR, and an IMU, enabling synchronous multi-source data collection. The raw data collected is stored and preprocessed after being aligned with a unified timestamp, including missing value imputation, outlier removal, and normalization operations, to ensure the consistency and fusibility of data from different modalities.

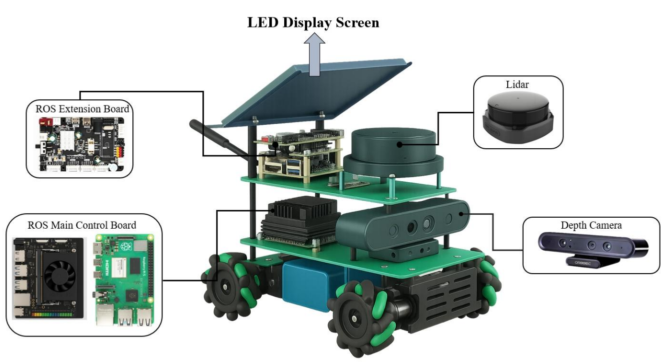


Fig. 2: AGV platform for the proposed system.

During operation, the Wi-Fi module collects RSSI vectors from surrounding access points (APs) to construct a fingerprint database, the LiDAR provides SLAM mapping and pose trajectory estimation, and the IMU makes high-frequency measurement of linear acceleration and angular velocity data. All data is uploaded to the edge computing unit for local processing and model prediction.

B. DNN-Based Wi-Fi RSSI Fingerprinting

To improve the localization accuracy of Wi-Fi RSSI fingerprinting in static areas, we use a DNN regression model to estimate the 2D coordinates of the AGV through non-linear mapping between a high-dimensional RSSI vector and physical space 2D coordinates under a supervised learning framework.

As shown in Fig. 3, the overall process of the Wi-Fi localization module is divided into four stages: data preparation, feature modeling, model training and prediction, and performance evaluation. Initially, the system loads a pre-built fingerprint database containing RSSI vectors collected at different reference points (RPs) and their corresponding 2D coordinates. To improve modeling quality and training stability, the original data undergoes preprocessing operations such as missing value filling, feature normalization, and outlier removal, and is then divided into training and test datasets.

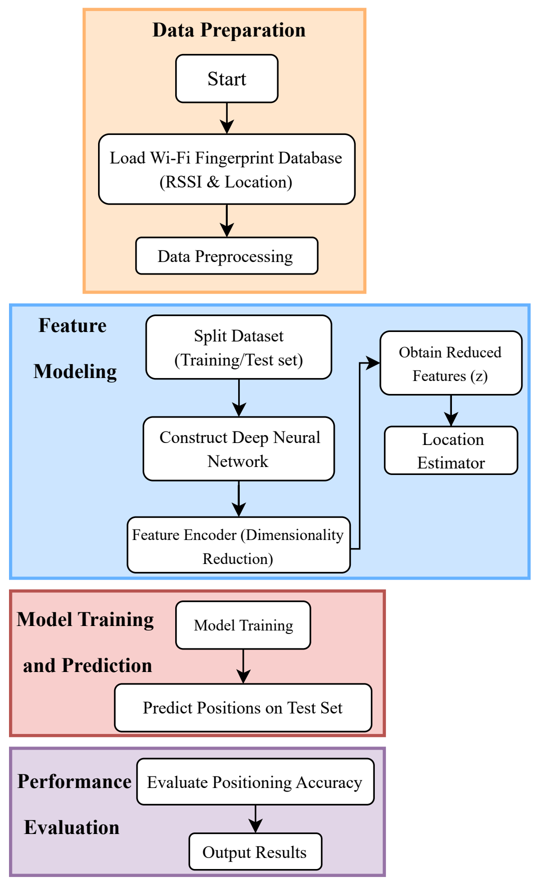


Fig. 3: Workflow of the Wi-Fi RSSI fingerprinting based on a DNN model.

A multi-layer feedforward neural network is used for the feature modeling stage to extract potential spatial distribution patterns from high-dimensional RSSI vectors. Table. I summarizes the parameter values of the DNN model. The network consists of seven linear transformation units, each with a rectified linear unit (ReLU) activation function and a dropout regularization mechanism with a dropout rate of 0.2

to suppress overfitting and improve generalization ability. The model input is 219-dimensional RSSI features (corresponding to the number of valid APs), and the output is 2D localization coordinates (x, y).

TABLE I: DNN parameter values for 2D-coordinate regression.

Layer	Type	Input Dimension	Output Dimension	Activation / Dropout
1	Linear	219	109	ReLU / Dropout(0.2)
2	Linear	109	73	ReLU / Dropout(0.2)
3	Linear	73	54	ReLU / Dropout(0.2)
4	Linear	54	109	ReLU / Dropout(0.2)
5	Linear	109	109	ReLU / Dropout(0.2)
6	Linear	109	109	ReLU / Dropout(0.2)
7	Linear	109	2	None

The DNN model is trained using mean squared error (MSE) as the loss function and the Adam optimizer for parameter learning. During the testing phase, the trained model predicts the 2D coordinates of given RSSI vectors, whose localization performance is evaluated based on the Euclidean distances between predicted and actual coordinates. The final results show that the DNN model has strong robustness and generalization capabilities and can be used as global localization prior information in the fusion phase to effectively support continuous trajectory estimation in subsequent multi-source fusion.

C. LiDAR/IMU-Based Dynamic Positioning

To improve the localization continuity and robustness of the system in complex dynamic environments, we present a LiDAR-based SLAM method enhanced by IMU data. Although the DNN-based Wi-Fi RSSI fingerprinting can provide high accuracy in static environments, its generalization ability decreases in scenarios with severe signal fluctuations, multipath propagation, serious obstructions, and high fingerprint database update costs, making it difficult to achieve stable and continuous location estimation.

To address the issues of the DNN-based Wi-Fi RSSI fingerprinting, we propose a SLAM method based on multi-sensor observations without prior maps, enabling real-time environment mapping and autonomous localization. The proposed method is widely used in mobile robots and indoor navigation systems, with the key idea of estimating the mobile platform's trajectory while dynamically updating the spatial representation of the environment, thereby achieving the joint inference of spatial and motion states.

The proposed method utilizes a Gmapping algorithm based on the Rao-Blackwellized particle filtering to integrate 2D laser radar scan data with wheel odometer information, thereby generating a closed-loop, consistent, high-resolution 2D environment map through a probabilistic grid mapping model, while outputting a high-precision pose trajectory estimate [24]. Compared with other SLAM methods, Gmapping has advantages such as, mature implementation, low computational resource overhead, and strong closed-loop detection capabilities, making it particularly suitable for real-time mapping and localization tasks in small and medium scale indoor scenes.

To further improve the system's dynamic response capability under challenging conditions such as rapid motion, sharp turns, and partial sensor occlusion, we introduce an IMU as an auxiliary sensor. A typical IMU equipped with three-axis accelerometer and gyroscope can measure three-dimensional (3D) linear acceleration $\boldsymbol{a}=[a_x\ a_y\ a_z]^\intercal$ and angular velocity $\boldsymbol{\omega}=[\omega_x\ \omega_y\ \omega_z]^\intercal$, which can be modeled as follows:

$$\begin{aligned} \boldsymbol{a} &= \boldsymbol{a}_{\text{true}} + \boldsymbol{b}_a + \boldsymbol{n}_a, \\ \boldsymbol{\omega} &= \boldsymbol{\omega}_{\text{true}} + \boldsymbol{b}_\omega + \boldsymbol{n}_\omega, \end{aligned} \tag{1}$$

where $\boldsymbol{a}_{\text{true}}$ and $\boldsymbol{\omega}_{\text{true}}$ denote the true values of 3D acceleration and angular velocity, $\boldsymbol{b}_a$ and $\boldsymbol{b}_\omega$ the measurement biases, and $\boldsymbol{n}_a$ and $\boldsymbol{n}_\omega$ the measurement noise, respectively. The measurement biases and noise in (1) are modeled and corrected during the fusion stage using the EKF to improve localization stability and accuracy.

Note that the high sampling frequency of the IMU can also effectively compensate for the shortcomings of LiDAR in short-term motion estimation, improving the localization continuity and robustness of the system in highly dynamic environments.

D. EKF-Based Multi-Sensor Fusion Based on Wi-Fi, LiDAR, and IMU

EKF is a basis of the proposed sensor fusion framework for indoor localization, which integrates the low-frequency global observations from Wi-Fi RSSI fingerprinting and the high-frequency motion data from an IMU. The objective is to produce accurate, continuous, and drift-resilient estimates of the platform's state in real time.

1) State Transition and Observation Models: The state vector and the control input of a system at time step k are defined as follows:

$$\begin{aligned} \boldsymbol{X}_k &= \begin{bmatrix} x_k & y_k & \theta_k & v_k & \omega_k \end{bmatrix}^\intercal, \\ \boldsymbol{u}_k &= \begin{bmatrix} a_k & \omega_k \end{bmatrix}^\intercal, \end{aligned} \tag{2}$$

where x_k and y_k represent the x and y coordinates of a 2D position, θ_k is the heading angle, v_k is the linear velocity, ω_k is the angular velocity, a_k is the measured linear acceleration, and ω_k is the measured angular velocity, respectively. The state vector $\boldsymbol{X}_k$ captures both kinematic and orientation dynamics for modeling the mobile platform behavior, while the control input $\boldsymbol{u}_k$ represents the measurement data from the IMU.

2) Prediction: Based on the prior state and control input, the predicted state $\hat{\boldsymbol{X}}_{k|k-1}$ can be obtained assuming constant acceleration:

$$\hat{\boldsymbol{X}}_{k|k-1} = \begin{bmatrix} x_{k-1} + v_{k-1}\Delta t \cos\theta_{k-1} \\ y_{k-1} + v_{k-1}\Delta t \sin\theta_{k-1} \\ \theta_{k-1} + \omega_{k-1}\Delta t \\ v_{k-1} + a_k \Delta t \\ \omega_k \end{bmatrix}, \tag{3}$$

where Δt denotes the sampling interval.

The corresponding covariance matrix is propagated using a linearized transition model:

$$P_{k|k-1} = F_k P_{k-1} F_k^{\intercal} + Q_k, \tag{4}$$

where F_k is the Jacobian matrix of the transition function concerning X_k, and Q_k is the process noise covariance capturing system model uncertainties.

3) Update: The predicted state $\hat{X}_{k|k-1}$ can be updated based on a newly-observed 2D location from Wi-Fi RSSI fingerprinting, which is modeled by

$$z_k = H_k X_k + v_k, \tag{5}$$

and

$$z_k = \begin{bmatrix} x_k^{\text{obs}} \\ y_k^{\text{obs}} \end{bmatrix}, \quad H_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad v_k \sim \mathcal{N}(0, R_k), \tag{6}$$

where z_k represents the newly-observed 2D location, H_k denotes the observation matrix that maps the full state vector X_k to the measured position components, and v_k is a zero-mean Gaussian noise with covariance matrix R_k, respectively.

The Kalman gain, which balances the trust between prediction and observation, is given by

$$K_k = P_{k|k-1} H_k^{\intercal} \left(H_k P_{k|k-1} H_k^{\intercal} + R_k \right)^{-1}. \tag{7}$$

Using the innovation term, the state can be updated as follows:

$$\hat{X}_k = \hat{X}_{k|k-1} + K_k \left(z_k - H_k \hat{X}_{k|k-1} \right). \tag{8}$$

The covariance matrix, too, is updated as follows:

$$P_k = (I - K_k H_k) P_{k|k-1}. \tag{9}$$

This EKF formulation supports the recursive fusion of asynchronous, heterogeneous sensor data. By leveraging the stability of Wi-Fi RSSI fingerprinting and the temporal resolution of the IMU signals, the proposed approach produces continuous and accurate indoor localization estimates that are robust to signal fluctuation, drift, and environmental noise.

IV. Experimental Results

We conduct experiments in the IR building at XJTLU, where we estimate the location of the AGV shown in Fig. 2 moving along the corridors of the 6th, 7th, and 8th floors. As the AVG moves along predefined trajectories, it collects data using the equipped Wi-Fi, LiDAR, and IMU modules.

For comparative analyses, we investigate the localization performance of the following three methods:

- DNN-based Wi-Fi RSSI fingerprinting (Wi-Fi).
- LiDAR/IMU-based dynamic positioning (LiDAR/IMU).
- EKF-based multi-sensor fusion based on Wi-Fi, LiDAR, and IMU (EKF).

As a performance metric, we use *mean 2D error*, which is defined over the test positions of a given trajectory as follows:

$$\text{Mean 2D Error} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\left(x_i^{\text{true}} - x_i^{\text{pred}}\right)^2 + \left(y_i^{\text{true}} - y_i^{\text{pred}}\right)^2}, \tag{10}$$

where N is the number of the test positions, and $(x_i^{\text{true}}, y_i^{\text{true}})$ and $(x_i^{\text{pred}}, y_i^{\text{pred}})$ are 2D coordinates of true and predicted positions, respectively.

A. Multi-Floor Trajectory Estimation

The floor plans of the 6th, 7th, and 8th floors of the IR building are similar to one another, whose example is shown in Fig. 4 for the experimental path for the AGV on the 7th floor. Repeating the same experiment on the three different floors, we obtained the three maps shown in Fig. 5 (a)–(c), each of which includes the three estimated trajectories for Wi-Fi, LiDAR/IMU, and EKF together with a true trajectory (i.e., ground truth). The mean 2D errors of the nine trajectories (i.e., three methods on three floors) are summarized in Table II.

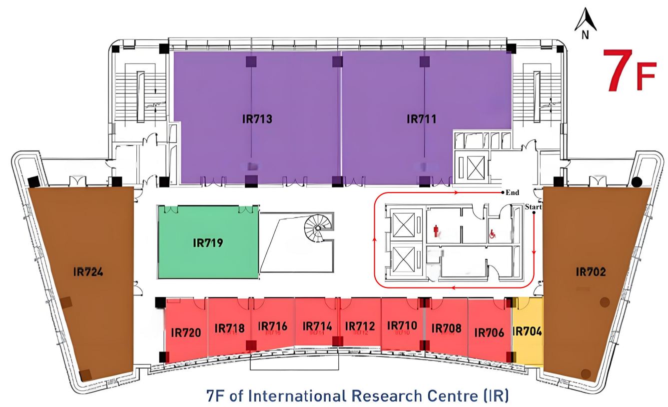
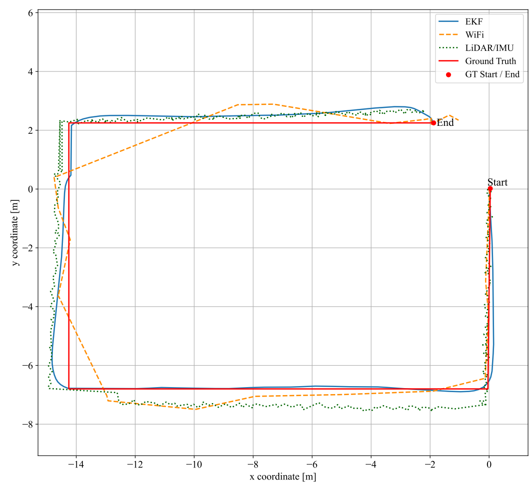


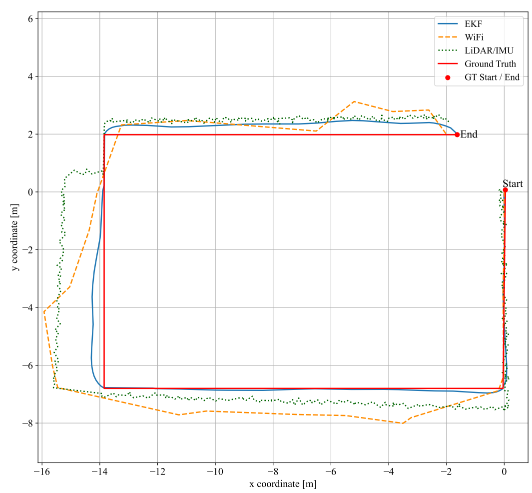
Fig. 4: Experimental path with its start and end points on the 7th floor of the IR building.

Fig. 5 shows that the estimated trajectories by the proposed EKF-based multi-sensor fusion are more stable and closer to the true trajectories in general than those by Wi-Fi and LiDAR/IMU, which is also confirmed by the mean 2D errors summarized in Table II. These experimental results demonstrate the effectiveness of the EKF-based multi-sensor fusion in addressing the issues of the drift and offset by Wi-Fi and LiDAR/IMU methods in indoor environments, which prevents them from reliably estimating the exact positions of a moving object.

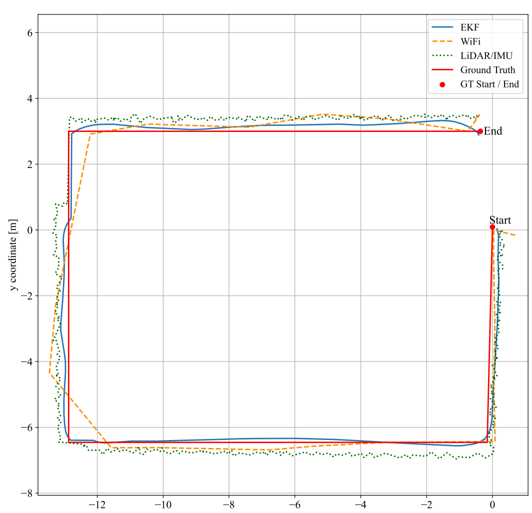
Taking the 7th floor as an example, the mean 2D errors of Wi-Fi and LiDAR/IMU are 0.8792 m and 0.8553 m, respectively, while that of EKF is only 0.2449 m. The results on the 6th and 8th floors also show similar trends, verifying that the multi-sensor fusion achieved by EKF can effectively reduce localization errors and improve the system's robustness and stability against different environmental disturbances.



(a) 6th Floor.



(b) 7th Floor.



(c) 8th Floor.

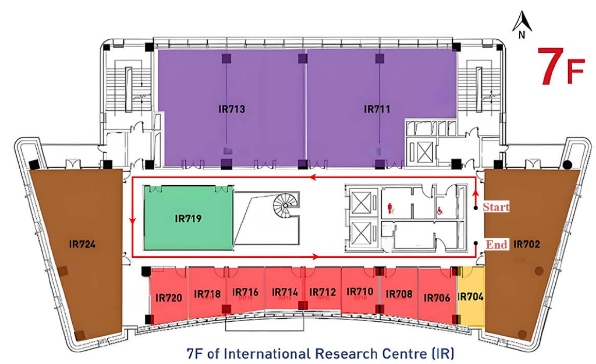
Fig. 5: Estimated trajectories on the floors of the IR building.

TABLE II: Comparison of the mean 2D errors over multiple floors.

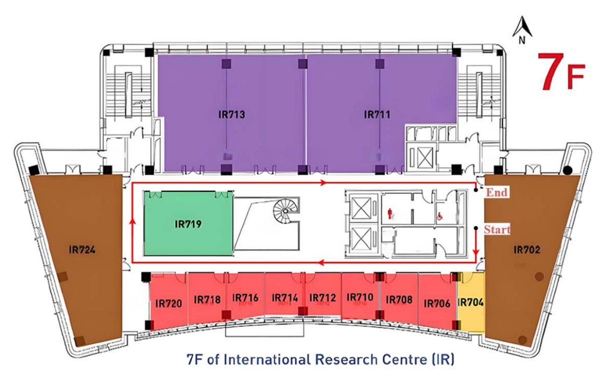
Floor	Method	Mean 2D Error [m]
6th	Wi-Fi	0.5127
	LiDAR/IMU	0.6233
	EKF	**0.3781**
7th	Wi-Fi	0.8792
	LiDAR/IMU	0.8553
	EKF	**0.2449**
8th	Wi-Fi	0.8624
	LiDAR/IMU	0.6756
	EKF	**0.2634**

B. Forward and Backward Trajectory Estimation

To further investigate the stability and adaptability of the implemented localization system under different path settings, we conducted experiments for forward and backward trajectory estimation. Fig. 6 (a) and (b) show the estimated trajectories by the three localization methods for the forward (clockwise) and backward (counterclockwise) paths on the 7th floor of the IR building, respectively. The mean 2D errors of the six trajectories (i.e., three methods for forward and backward paths) are summarized in Table III.



(a) Forward path.



(b) Backward path.

Fig. 6: Forward and backward paths with their start and end points on the 7th floor of the IR building.

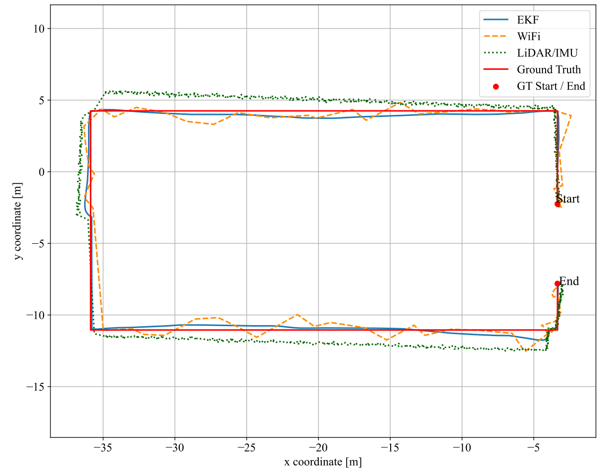
Like the results of the multi-floor trajectory estimation presented in Section IV-A, Fig. 6 shows that the estimated trajectories, by the proposed EKF-based multi-sensor fusion, are more stable and closer to the true trajectories in general than those by Wi-Fi and LiDAR/IMU under the forward and backward path scenarios, which is again confirmed by the mean 2D errors summarized in Table III. Though the system faces the differences in motion patterns and the changes in observation distribution when moving along the backward path, the EKF-based multi-sensor fusion can still maintain a relatively lower estimation deviation, demonstrating its strong model generalization capabilities.

Compared to EKF, Wi-Fi and LiDAR/IMU show significantly higher estimation errors, especially right after the turns on the three corners, due to their susceptibility to signal obstruction and inertial drift, which leads to cumulative estimation bias; in the case of the proposed multi-sensor fusion method, because EKF uses a prediction-correction mechanism to effectively fuse global observations and local dynamics, it can successfully suppress cumulative errors and thereby improve overall localization accuracy and robustness.

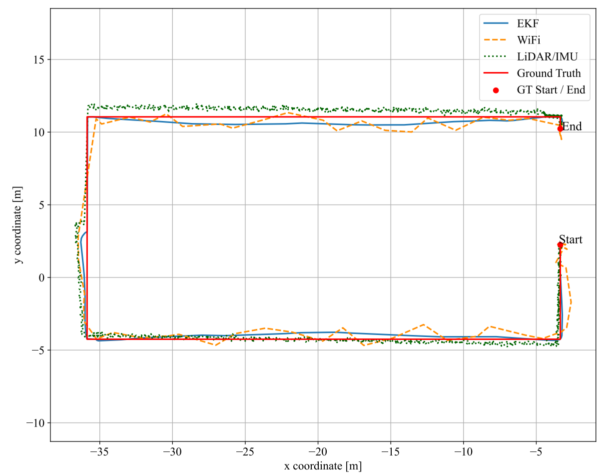
In addition, the comparison of the results for the forward and backward paths reveals that the system exhibits good directional symmetry and path invariance. These experimental results further demonstrate the capability of the proposed EKF-based multi-sensor fusion method in handling different path configurations and dynamic disturbances in real-world deployments.

TABLE III: Comparison of the mean 2D errors over forward and backward paths.

Path	Method	Mean 2D Error [m]
Forward	Wi-Fi	1.3404
	LiDAR/IMU	2.8803
	EKF	**0.3672**
Backward	Wi-Fi	1.2651
	LiDAR/IMU	1.8968
	EKF	**0.2851**



(a) Forward path.



(b) Backward path.

Fig. 7: Estimated trajectories for the forward and backward paths on the 7th floor of the IR building.

V. Conclusions

We have proposed an indoor localization and navigation framework that integrates DNN-based Wi-Fi RSSI fingerprinting, IMU-assisted dynamic localization, and LiDAR environmental perception capabilities through multi-sensor fusion based on the EKF algorithm. We implemented the proposed localization system on an AGV and conducted comprehensive field experiments with it on the 6th, 7th, and 8th floors of the IR building at XJTLU to verify the stability and localization accuracy of the proposed EKF-based multi-sensor fusion method under multipath interference and signal drift.

The experimental results demonstrate that the multi-sensor fusion method based on the EKF maintains high positioning accuracy under various path conditions and floor environments, with the mean 2D error consistently controlled between 0.2449 m and 0.3781 m. In contrast, the DNN-based Wi-Fi regression model exhibits significant error fluctuations, with errors ranging from 0.5127 m to 1.3404 m under signal obstruction or multipath effects. The LiDAR/IMU-based dynamic positioning, too, exhibits greater positioning errors distributed between 0.6233 m and 2.8803 m due to the IMU cumulative error issues. By fusing multi-source sensor data, the EKF effectively suppresses the deviations and uncertainties caused by the aforementioned error sources, making the overall trajectory more closely aligned with the actual path. This significantly enhances the system's robustness and adaptability in complex environments.

Acknowledgment

This work was supported in part by Xi'an Jiaotong-Liverpool University (XJTLU) Summer Undergraduate Research Fellowships (under Grant SURF-2024-0311).

REFERENCES

[1] N. H. A. Wahab, N. Sunar, S. H. Ariffin, K. Y. Wong, Y. Aun *et al.*, "Indoor positioning system: A review," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 6, 2022.

[2] X. Guo, N. Ansari, F. Hu, Y. Shao, N. R. Elikplim, and L. Li, "A survey on fusion-based indoor positioning," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 566–594, 2019.

[3] F. Gustafsson and F. Gunnarsson, "Positioning using time-difference of arrival measurements," in *Proc. ICASSP'03*, vol. 6. IEEE, 2003, pp. VI–553.

[4] S. Yiu, M. Dashti, H. Claussen, and F. Perez-Cruz, "Wireless RSSI fingerprinting localization," *Signal Processing*, vol. 131, pp. 235–244, 2017.

[5] H. Yan, Q. Shan, and Y. Furukawa, "RIDI: Robust IMU double integration," in *Proc. ECCV 2018*, 2018, pp. 621–636.

[6] R. Jirawimut, P. Ptasinski, V. Garaj, F. Cecelja, and W. Balachandran, "A method for dead reckoning parameter correction in pedestrian navigation system," *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 1, pp. 209–215, 2003.

[7] F. Caron, E. Duflos, D. Pomorski, and P. Vanheeghe, "GPS/IMU data fusion using multisensor Kalman filtering: Introduction of contextual aspects," *Information Fusion*, vol. 7, no. 2, pp. 221–230, 2006.

[8] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "IMU-based gait recognition using convolutional neural networks and multi-sensor fusion," *Sensors*, vol. 17, no. 12, p. 2735, 2017.

[9] Y. Wu *et al.*, "Indoor localization system based on IMU and Wifi," *Sensors*, vol. 15, no. 7, pp. 17 947–17 960, 2015.

[10] G. He, X. Yuan, Y. Zhuang, and H. Hu, "An integrated GNSS/LiDAR-SLAM pose estimation framework for large-scale map building in partially GNSS-denied environments," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2020.

[11] K. Jo, K. Chu, and M. Sunwoo, "Interacting multiple model filter-based sensor fusion of GPS with in-vehicle sensors for real-time vehicle positioning," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 329–343, 2011.

[12] M. I. Ribeiro, "Kalman and extended Kalman filters: Concept, derivation and properties," *Institute for Systems and Robotics*, vol. 43, no. 46, pp. 3736–3741, 2004.

[13] G. A. Einicke and L. B. White, "Robust extended Kalman filtering," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2596–2599, 2002.

[14] X. Liu, B. Zhou, P. Huang, W. Xue, Q. Li, J. Zhu, and L. Qiu, "Kalman filter-based data fusion of Wi-Fi RTT and PDR for indoor localization," *IEEE Sensors J.*, vol. 21, no. 6, pp. 8479–8490, 2021.

[15] D. Feng, C. Wang, C. He, Y. Zhuang, and X.-G. Xia, "Kalman-filter-based integration of IMU and UWB for high-accuracy indoor positioning and navigation," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3133–3146, 2020.

[16] A. Barrau, "Non-linear state error based extended Kalman filters with applications to navigation," Ph.D. dissertation, Mines Paristech, 2015.

[17] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola, and C. Fischione, "A survey of enabling technologies for network localization, tracking, and navigation," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3607–3644, 2018.

[18] P. Zhou, H. Wang, R. Gravina, and F. Sun, "WIO-EKF: Extended Kalman filtering-based Wi-Fi and inertial odometry fusion method for indoor localization," *IEEE Internet Things J.*, vol. 11, no. 13, pp. 23 592–23 603, 2024.

[19] Q. Tang, J. Liang, and F. Zhu, "A comparative review on multi-modal sensors fusion based on deep learning," *Signal Processing*, vol. 213, p. 109165, 2023.

[20] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2568–2599, 2019.

[21] K. Liu and et al., "Fusion of Wi-Fi, inertial and magnetic sensors for indoor localization using deep learning," *IEEE Sensors J.*, vol. 19, no. 13, pp. 4821–4830, 2019.

[22] Y. Gu and et al., "UWB and LiDAR data fusion for indoor mobile robot localization," *Sensors*, vol. 20, no. 8, p. 2345, 2020.

[23] H. Zhou *et al.*, "A deep learning framework for robust indoor localization using LSTM and multi-sensor fusion," *Sensors*, vol. 22, no. 4, p. 1654, 2022.

[24] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with Rao-Blackwellized particle filters," *IEEE Transactions on Robotics*, vol. 23, no. 1, pp. 34–46, 2007.