# A real-world case study: Smartphone-based Indoor Localization

Ming Yuan Kee
*College of Computing and Data Science (CCDS)*
*Nanyang Technological University*
Singapore, Singapore
mkee004@e.ntu.edu.sg

Sathananthar Suresh
*College of Computing and Data Science (CCDS)*
*Nanyang Technological University*
Singapore, Singapore
suresh011@e.ntu.edu.sg

Robert George Susanto
*College of Computing and Data Science (CCDS)*
*Nanyang Technological University*
Singapore, Singapore
robe0006@e.ntu.edu.sg

Sui Jinhong
*College of Computing and Data Science (CCDS)*
*Nanyang Technological University*
Singapore, Singapore
suij0003@e.ntu.edu.sg

***Abstract*— Even with great technological advances, indoor localisation faces numerous challenges, impeding the accurate determination of a person's location within indoor environments. Using the data provided by the Kaggle competition: Indoor Location & Navigation, this report aims to create a model to enhance location accuracy in indoor settings to improve indoor navigation and tracking systems. The data, namely the indoor signatures of geomagnetic fields, WiFi, and iBeacon, and the ground truth location of two sites, will be pre-processed and analyzed before building a deep learning model for location prediction.**

***Keywords—Indoor Localization, Simultaneous Localization And Mapping (SLAM), Deep Learning, Geomagnetic field strength, iBeacon signal, WiFi signal, ground truth location***

## I. Introduction

In contrast to outdoor localization, which relies on the relatively precise Global Positioning System (GPS), indoor localization faces numerous challenges. It is challenging for GPS to be used for Indoor localization as the signal from the satellite is not strong enough to penetrate building walls to capture the signal. Various technologies have been developed for indoor localization, yet each comes with its own set of obstacles. For instance, some methods, such as device-free techniques, require extensive infrastructure deployment indoors, leading to significant implementation costs while device-free methods, such as using surveillance cameras to monitor individuals' locations, raise privacy concerns regarding individuals captured on camera. Furthermore, these surveillance cameras may not consistently provide accurate location data, as their view can be obstructed.

Other Indoor Localization methods, such as Surround Sense, utilizes a multi-dimensional fingerprinting approach that captures ambient sound, light intensity, user movement, and Wi-Fi SSIDs at various points within a space to ascertain an individual's location. These methods demand many resources due to the manual collection of blanket data.

Powerline Simultaneous Localization And Mapping (SLAM) is an Indoor Localization method where electromagnetic frequency (EMF) sensors are used to sense the electromagnetic radiation produced by the powerline at different indoor locations. Combining EMF readings, which are not affected by the user's height, geomagnetic field readings, and electrical appliances, can achieve good indoor location accuracy. Nevertheless, this is not a commercialized indoor localization method.

This report aims to develop a fingerprinting model capable of identifying an individual's location within an indoor environment using real-time sensor data. The data provided will be pre-processed and visualized through graphical representations using matplotlib and mpld3. The proposed model is a deep learning approach utilising a multi-layer perceptron (MLP) architecture, achieving a Mean Absolute Error (MAE) of 36.72 metres.

## II. Dataset

The actual dataset from the Kaggle competition, Indoor Location & Navigation, includes data from 204 sites in China. Only data from two sites will be analyzed based on the source code provided [1], namely site 1 of ID 5dd3d7732a57a34356595932 and site 2 of ID 5dbc1d84c1eb61796cf7c010. The data from both sites consists of indoor sensing information collected across all floors of each site, where surveyors walked to various paths and stopped a few locations of the site to collect data.

TABLE I. INFORMATION USED FOR EACH SENSING DATA

| **Sensing Data** | **Values Used** |
|---|---|
| Ground truth location | 2D (x, y) |
| Accelerometer | 3D (x, y, z) |
| Gyroscope | 3D (x, y, z) |
| Rotation Vector | 3D (x, y, z) |
| Magnetic field strength | 3D (x, y, z) |
| WiFi | SSID, BSSID, RSSI |
| iBeacon | UUID, Major ID, Minor ID, RSSI |

The table above provides the ground truth location of each position of the path the surveyor walked. IMU data such as accelerometer, gyroscope and rotation vector are also given. At each location, WiFi and iBeacon signals, as well as geomagnetic field strength (GFS), are collected. Although uncalibrated data from the accelerometer, gyroscope and GFS were given, they were not used in this report. For the WiFi data, SSID is the name of the WiFi signal, BSSID is the unique MAC address, and RSSI is the signal strength received. For iBeacon, UUID is the unique identifier of each beacon used

by the same organization. UUID, Major ID, and Minor ID identify a beacon uniquely. Similar to a WiFi signal, RSSI is the strength of the beacon's received signal.

The metadata includes PNG files for each site's floor, specifying the height and width of the floor map images and GeoJSON files representing geographical features on the floor maps.

The spatiotemporal data is pre-processed and transformed from its raw format into meaningful information. The surveyor's position is estimated using IMU data, including acceleration and orientation, for geomagnetic field strength, WiFi, and iBeacon signals. Using the acceleration data collected, the surveyor took a step at a certain timestamp when there was a peak in acceleration. Using the acceleration value range, the length of the step taken by the surveyor is calculated. To know in which direction the surveyor took that step, AHRS information is used. Using the waypoint locations, which are accurately determined, the estimated positions of the surveyor can be determined with the movement information calculated.

The preprocessing approach for Wi-Fi, iBeacon, and magnetic data follows a consistent method. For each type of sensor data, relevant readings—such as Wi-Fi RSSI, iBeacon RSS values, or magnetic field strength, the timestamps of the surveyor's estimated locations are matched with the timestamp of each sensor data type. The corresponding sensor data type readings with the closest timestamp to the surveyor's location are selected and paired accordingly for each estimated location. If a WiFi signal given by a unique BSSID is detected more than once at the same estimated position, the average WiFi signal strength will represent that BSSID at that location. This gives a more accurate representation of the WiFi signal strength. The same applies to iBeacon, where a unique UUID, Major ID, and Minor ID are given. The preprocessed data can then be used to generate heatmaps, enabling the visualization of spatial patterns in signal strength or magnetic anomalies, which are essential for indoor localization.

## III. Essential Task 1: Visualization of Waypoints

Looking at the figure below, it is obvious that plotting waypoints of all floors together does not give much information. This is because different floors might have different layouts, and the above graph does not explain why the waypoints are in a certain shape. Hence, the waypoints will be evaluated on each floor of the site.

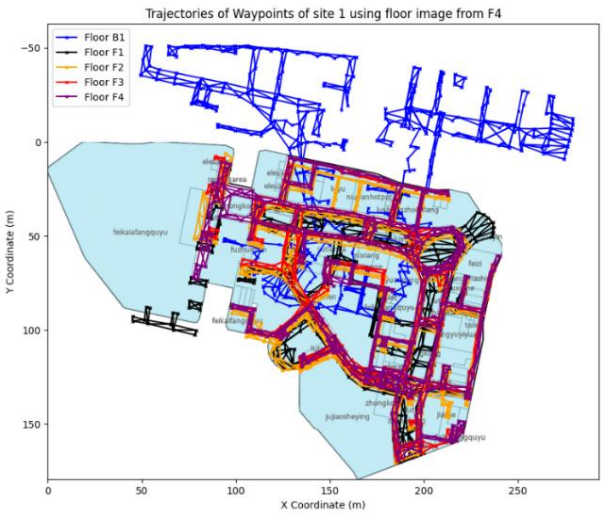



Fig. 1.1. Waypoints of all floors in site 1

The surveyor has traversed multiple distinct trajectories for each floor, each characterized by its own starting and ending points, as seen in the figure below. The below path contains 18 locations.


Fig. 1.2. One of the many paths walked by the surveyor in F1 of site 1

All surveyor trajectories are flipped over the y-axis before being overlaid on the corresponding floor map image provided in the metadata file. The figure below shows the waypoint location of all trajectories covered by the surveyors on floor B1 of site 1. In most trajectories, the user went in a zig-zag trajectory when collecting data to cover both the left and right sides of the corridors. We see that certain floor areas are not covered by any trajectories at all. This could be that those areas are out of bounds to the surveyors. Hence, sensing data cannot be collected. This could serve a limited class coverage problem where the model built from such training data might not generalize well to locations of B1, which is not covered in the training data.

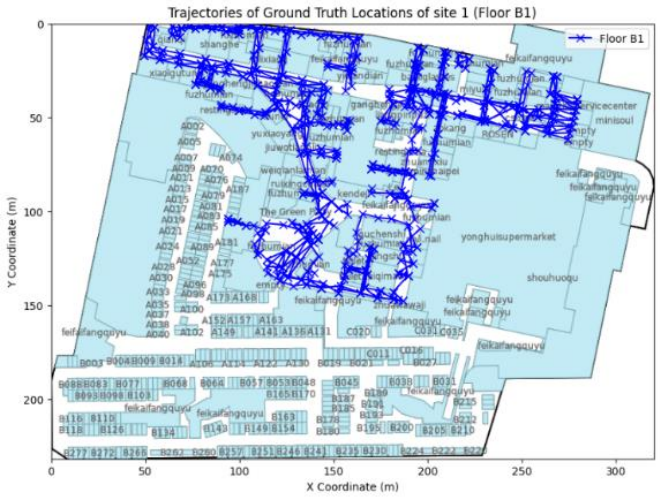



Fig. 1.3. All waypoint trajectories of B1 of site 1

## IV. Essential Task 2: Visualize Geomagnetic Heatmap

The geomagnetic field is a natural phenomenon produced by the Earth's core, and it can be detected by the magnetometer sensor embedded in most mobile devices. Mobile devices can use this data to sense variations in magnetic field strength, which may occur due to nearby metallic objects or structural anomalies. These variations in the geomagnetic field are helpful for indoor localization, as they can help map unique magnetic signatures of specific indoor areas, aiding in navigation when GPS signals are weak or unavailable.

### *Approach Description*

For this task, we aimed to visualize the geomagnetic field strength across the floor plan of Site 1, specifically on Level F1. The objective was to represent the magnetic field intensities at different locations within the floor area. This can be useful for indoor localization systems where geomagnetic anomalies are key features. We began by preprocessing the data, where geomagnetic readings were extracted from the sensor dataset. Each data point contained magnetic field values in the x, y, and z directions. The magnetic field strength, measured in microteslas (μT), was calculated using the magnitude of the magnetic vector as $\sqrt{x^2 + y^2 + z^2}$. The corresponding positions (x, y coordinates) where the recorded geomagnetic data were also extracted for visualization.

Next, we generated a heatmap to visualize the magnetic field distribution. The extracted positions and magnetic strength values were plotted over the floor plan. Each point on the scatter plot represented a recorded position, and the point's colour indicated the intensity of the geomagnetic field at that location. The floor plan was superimposed as the plot's background to provide context. A colour scale, ranging from blue (lower magnetic strength) to red (higher magnetic strength), was used to represent the field intensities visually, and a colour bar was included to make interpreting the data easier. The marker size was dynamically adjusted based on the floor plan dimensions to ensure visibility, and the heatmap was scaled to match the building's aspect ratio for accurate spatial representation.

### *Result and Presentation*

The resulting heatmap revealed the magnetic field distribution on Site 1, Floor F1 (Fig 2.1). The visualization shows that the magnetic field strength is not uniformly distributed across the floor.

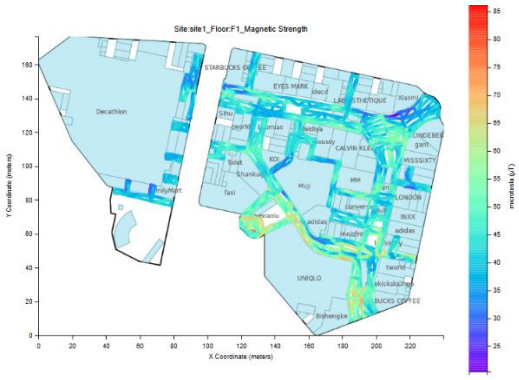



Fig. 2.1. Magnetic Strength Heatmap for Site 1, Floor F1.

Some regions, highlighted in red, exhibit significantly higher magnetic intensity, which may be caused by nearby electrical systems, metal structures, or other factors that disturb the Earth's magnetic field. These high-intensity areas could be potential points of interest for further investigation, especially for indoor positioning systems that leverage magnetic anomalies. In contrast, lower-intensity areas are marked in blue or purple. The variability in magnetic strength across the floor may also reflect structural differences, such as the presence of steel reinforcement in walls or floors or proximity to electrical equipment.

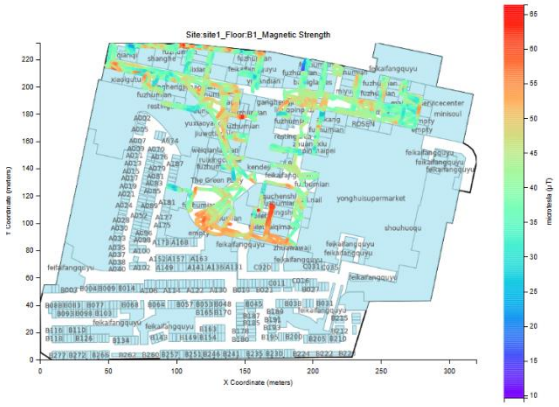



Fig. 2.2. Magnetic Strength Heatmap for Site 1, Basement B1.

In the provided magnetic strength heatmaps for Site 1 Floor F1 (Fig. 2.1) and Basement B1 (Fig. 2.2), we observe notable variations in magnetic intensity across the two floors. In the basement (B1), several regions with higher magnetic strength, particularly near certain stores and structural elements, are likely due to reinforced metal structures or proximity to electrical equipment, which can influence the magnetic field. On Floor F1, the magnetic strength is more evenly distributed, although the high-intensity areas still exist, particularly along certain pathways and near commercial units. These variations in magnetic field strength are critical for indoor localization systems, as they can provide distinguishable magnetic signatures or anomalies at specific locations. Using these unique magnetic signatures as reference points, mobile devices can better estimate their location within the building when combined with other sensor data, such as

Wi-Fi or accelerometer readings, thus improving indoor positioning accuracy.

In conclusion, the geomagnetic heatmaps for Site 1's Floor F1 and Basement B1 highlight the spatial variability of magnetic field intensities within the building. The magnetic field strength, influenced by structural elements such as metal reinforcements or electrical equipment, varies significantly across both floors, offering unique geomagnetic signatures. These variations are particularly useful for indoor localization systems, as mobile devices can leverage these distinct magnetic anomalies to improve location accuracy in environments where GPS signals are weak or absent. The visualizations presented provide valuable insights into the magnetic environment of the building, supporting potential applications in indoor navigation and positioning.

## V. Essential Task 3: Visualize Wi-Fi RSS Heatmap of 3 Wi-Fi APs

A Wi-Fi Access Point (AP) is a network device that enables devices to connect to a wireless local area network. Devices within the AP's coverage area can detect signals from the AP, even if they are not actively connected to the Wi-Fi. The received signal strength (RSS) refers to the power of a signal detected at the receiver's antenna and is influenced by factors such as (1) Transmission power: Higher power results in a stronger RSS. (2) Distance between the transmitter and receiver: Devices closer to the transmitter receive stronger signals. (3) Radio environment: Obstructions like metal, walls, and water can weaken the RSS. As such, RSS is commonly used for localization, as it serves as a proxy for estimating the distance between a device and the AP. RSS is measured on a scale from 0 to -100 dBm, where 0 represents the strongest signal and -100 is the weakest.

In this task, we visualized floor F1 from site 1 to identify the best proxy for coverage. We hope the insights can guide us in engineering additional features when tackling the bonus prediction task. In addition, a larger coverage area reduces the need for multiple APs, lowering capital costs and minimizing the amount of data required for precise localization, which speeds up processing and real-time localization. We identified 3 APs based on the following criteria: (1) The AP with the most connected devices. (2) The AP with the largest Manhattan distance between devices receiving the highest and lowest RSS. (3) The AP with the largest Manhattan distance between devices receiving the highest and lowest RSS.

In the following plots, each dot represents a single device. Devices with stronger RSS values are shown in red, while weaker RSS values are in purple. The red crosses mark the estimated locations of the Wi-Fi APs based on the received RSS.

In the following plots, signal strength is typically stronger when devices are close to the predicted AP location, especially in areas without physical barriers, such as shops. However, the RSS is significantly weakened when shops between the AP and the devices are present. This trend was most evident in regions with poor coverage, where many shops were located between the AP and those areas.

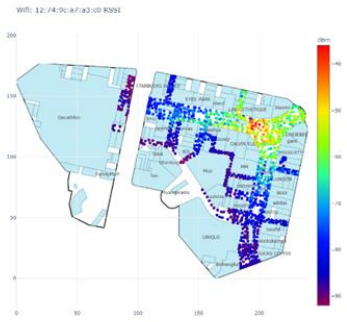


Fig. 3.1. Heatmap of AP with the highest number of connected devices

In the above plot, we anticipated that the AP with the most connected devices would be a good coverage indicator, as people tend to move around rather than stay in a single area. A higher number of connected devices generally suggests broader coverage. This was validated by the plot, which showed extensive coverage across most of the map, with weaker coverage in areas where many shops were situated between the AP and those regions. We predicted that the AP's location is near the cluster with high RSS values.

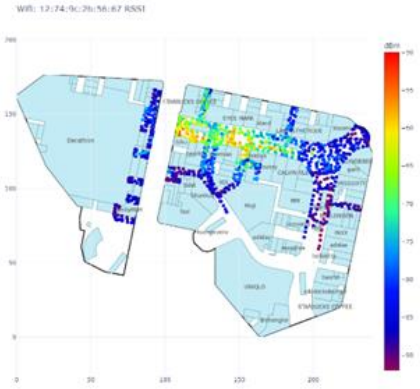


Fig. 3.2. Heatmap of AP with the highest Manhattan distance between the device with the highest and lowest RSS

In the plot above, we are unsure about the location of the AP due to 2 distinct clusters of high RSS values. We hypothesized that a greater distance between the devices with the highest and lowest RSS would indicate better coverage. This is because broader coverage typically results in a larger gap between the device closest to the AP (with the strongest RSS) and the one furthest away (with the weakest RSS). The plot demonstrated decent coverage, though not as strong as the previous one. The larger Manhattan distance could indicate that the path is not direct, and shops between devices and the AP likely weakened the RSS.

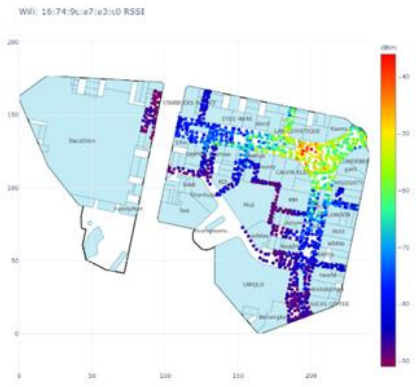


Fig. 3.3. Heatmap of AP with the lowest total RSS

In the plot above, we anticipated the AP to be located in a similar position as in the first plot, close to the cluster with a high concentration of RSS values in that area. We hypothesized that lower total RSS scores would correspond to wider coverage, as broader coverage suggests more connected devices, increasing the presence of low RSS values and reducing the total RSS score. This was confirmed, as the AP with the lowest total RSS had almost identical coverage as the AP with the highest number of connected devices. This choice also served as a good proxy, as the coverage area was nearly identical to that in the first plot.

## VI. ESSENTIAL TASK 4: VISUALIZE IBEACON HEATMAP

An iBeacon is a location-based technology created by Apple that utilizes low-energy Bluetooth (BLE) transmitters to broadcast signals to mobile devices within proximity, for example our smartphones [2]. When a device captures the broadcast signal, it can trigger predefined actions, including displaying alerts or delivering contextually relevant content based on the distance to the beacon. In essence, iBeacon enables mobile devices to recognize their spatial vicinity to specific locations or objects and automate responses accordingly. This technology is widely implemented in sectors such as retail, where it enhances user experiences by providing personalized promotions and offers, streamlining navigation, and enabling contactless payment systems.

To assess the performance of iBeacon, we applied the following criteria: (1) The iBeacon with the highest number of detected interactions across all devices. (2) The iBeacon that exhibits the largest spatial distribution, as measured by the distance between devices recording the highest and lowest RSSI values. (3) The iBeacon with the broadest signal range, determined by the distance between the strongest and weakest signal points recorded. These metrics ensure optimal signal coverage and more efficient localization across the entire floor.

According to the iBeacon Manual book provided by Apple, a basic iBeacon entry consists of:

| Field | Size | Description |
|---|---|---|
| UUID | 16 bytes | Application developers should define a UUID specific to their app and deployment use case. |
| Major | 2 bytes | Further specifies a specific iBeacon and use case. For example, this could define a sub-region within a larger region defined by the UUID. |
| Minor | 2 bytes | Allows further subdivision of region or use case, specified by the application developer. |

Fig. 4.0 Fields of iBeacon provided by Apple [3]

To visualize the iBeacon, we will use the F1 floor from Site 1. For this example, I will be taking three iBeacon IDs to compare the differences between the three:

1. 9195B3AD-A9D0-4500-85FF-9FB0F65A5201_0_0
2. 6B76E28A-6FA2-48C9-8502-C1DAA388AB2C_47237_15305
3. E7FC9D3C-EF01-4B70-B280-2CF6D50FA5CA_13394_63898

Fig. 4.1. Heatmap of iBeacon signal strength of iBeacon: 9195B3AD-A9D0-4500-85FF-9FB0F65A5201_0_0

The plot above illustrates the distribution of signal strength (RSSI) for a particular iBeacon of ID 9195B3AD. The heatmap effectively visualizes the spatial coverage of the iBeacon signals across the entire floor plan, with warmer colors representing stronger signal strength (closer to -55 dBm) and cooler colors indicating weaker signals (closer to -100 dBm). iBeacons with higher signal strengths are observed to be in high-traffic areas, particularly near the entrances of prominent stores and the hallways. This suggests that these zones experience a higher number of signal interactions, likely due to increased pedestrian activity. As distance increases from these high-density regions towards areas with lower foot traffic or as users exit the stores, there is a noticeable decrease in signal strength, represented by cooler colors in the heatmap.

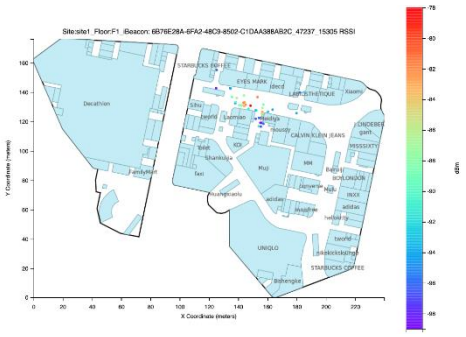


Fig. 4.2. Heatmap of iBeacon signal strength of iBeacon: 6B76E28A-6FA2-48C9-8502-C1DAA388AB2C_47237_15305

The plot for Figure 4.2 corresponding to iBeacon ID 6B76E28A demonstrates a more limited coverage area than Figure 4.1. Unlike the broader distribution observed previously, the signal is concentrated in a particular region and progressively weakens as distance from the source increases. This speculation is likely due to the greater distance from the iBeacon, while potential physical obstructions are also considered, which further diminish the strength through interference or absorption.

In contrast, the strongest signal is detected in a localized area just outside a certain store, suggesting the possible physical location of the planted iBeacon. This pattern is consistent with expected Bluetooth signal behaviour, where

signal strength is the highest near the iBeacon and decreases as distance increases.

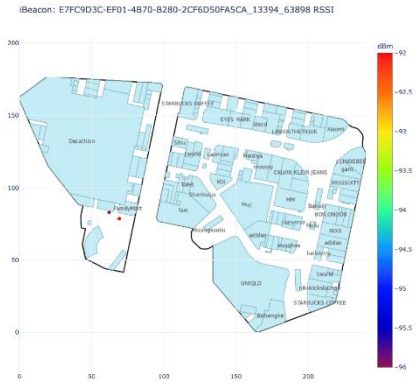



Fig. 4.3. Heatmap of iBeacon signal strength of iBeacon: E7FC9D3C-EF01-4B70-B280-2CF6D50FA5CA_13394_63898

The plot in Figure 4.3, corresponding to iBeacon ID E7FC9D3C, displays even more limited coverage compared to Figure 4.2. There are only two recorded RSSI values—one indicating the strongest signal (represented by red) and the other showing the weakest signal (represented by purple). This sharp contrast in signal strength, occurring within a small spatial area, suggests that that particular iBeacon's coverage is extremely constrained. The close proximity of these two signals, despite their iBeacon signal strength at opposite ends of the RSSI spectrum demonstrates the limited effective range of this iBeacon.

## VII. BONUS TASK 1: BUILD A DEEP LEARNING-BASED LOCATION FINGERPRINT MODEL

In this task, we developed a deep learning-based location fingerprint model to predict a user's location using a dataset containing fingerprint data from iBeacon, WiFi, and geomagnetic field strengths (GFS). The objective was to compare two models and evaluate whether including external reference points (iBeacon, WiFi) alongside motion sensor data (accelerometer, gyroscope, and rotational data) would yield better results than using only the data from the naturally occurring GFS.

Two models were built from the dataset from Site 1, Floor F1. **Model 1** integrates signal strength data from iBeacon, WiFi, GFS data, and motion sensor data from the accelerometer, gyroscope, and rotational sensors. The mean signal strength and sensor readings at each waypoint were used as features. **Model 2** relies solely on the median of GFS, which is naturally available at all parts of the indoor environment and does not rely on additional infrastructure.

In **Model 1**, the mean of multiple signal measurements at each waypoint was aggregated to create stable features. In contrast, **Model 2** focused on the median of GFS to minimize noise from potential outliers, such as interference from metallic objects. Both models used a multi-layer perceptron (MLP) architecture: Model 1 featured four fully connected layers, and Model 2 used three. Dropout and batch normalization were applied after each hidden layer to prevent overfitting. Dropout (0.3) and L2 regularization were used to generalize model performances across varied environments. The learning rate of 0.001 ran on 1000 epochs, and the Adam optimizer was used for both models. More information on model specifications is in the appendix.

The performance of both models was assessed using the mean absolute error (MAE) metric, with results displayed in Table II below. MAE is the absolute difference in distance of the predicted location from the model and the ground truth location provided. Figures 5.1 and 5.2 show boxplots comparing the MAE values for Models 1 and 2, respectively.

TABLE II. MODEL PERFORMANCE COMPARISONS

| Model | MAE (m) | Standard Deviation (m) |
|---|---|---|
| Model 1 | 45.73 | 25.56 |
| Model 2 | 36.72 | 22.22 |

The results indicate that **Model 2**, which relies solely on geomagnetic data, performs better than **Model 1**. **Model 2** achieved an MAE of 36.72 metres (m), while **Model 1** had a higher MAE of 45.73m. The lower standard deviation in Model 2 (22.22m) reflects more consistent performance, suggesting that GFS data is more stable and less influenced by external fluctuations than iBeacon and WiFi signals. **Figure 1** and **Figure 2** provide further evidence, with Model 2 showing a lower median distance difference (30 metres) than Model 1 (40 metres).

The results demonstrate that Model 2, which uses only geomagnetic field strength (GFS), outperforms Model 1 regarding prediction accuracy, with a mean error of 36.72 metres. This could be because the GFS data is collected simultaneously as the IMU data. In contrast, the WiFi and iBeacon signal data is collected with a few minutes of time difference from the time when the IMU data is collected. Hence, the GFS data is more accurate than the WiFi and iBeacon signal strength data. The reliance on naturally occurring magnetic fields offers advantages, such as reducing dependence on external infrastructure and yielding stable predictions across different indoor environments.

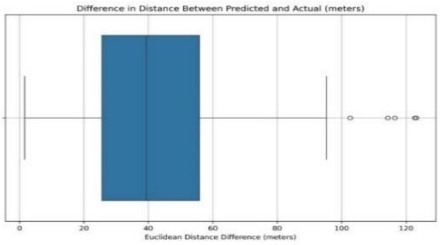



Fig. 5.1. Boxplot of difference in distance between actual and predicted location using model 1.

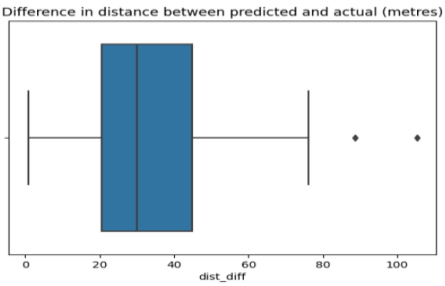



Fig. 5.2. Boxplot of difference in distance between actual and predicted location using model 2.

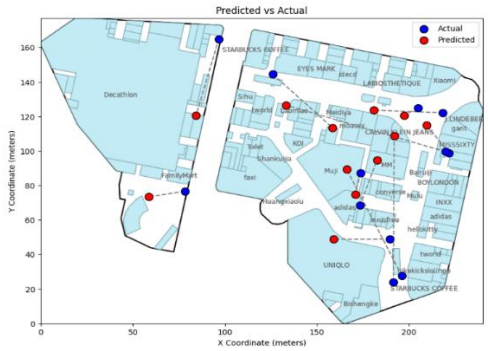



Fig. 5.3. Visualization of the predicted and actual points of model 2.

Figure 5.3 shows that while Model 2 performed better overall, it still has room for improvement. Although some predictions are close to the actual locations, others are significantly off, resulting in an overall MAE that is still not accurate enough for precise indoor localization.

Future work will focus on enhancing the fusion of motion sensor data (accelerometer, gyroscope, and rotational) with GFS data using advanced techniques like Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks to better capture the temporal dependencies in sensor data. Transfer learning methods will also be explored, enabling pre-trained models to adapt to new environments with minimal retraining. These advancements aim to further improve the robustness and accuracy of indoor localization systems, especially in GPS-denied environments.

## VIII. CONCLUSION

This report outlines developing and evaluating deep learning models for indoor localization using data from the Kaggle competition: Indoor Location & Navigation. After pre-processing the provided data, visualizations for ground truth locations, geomagnetic field strength (GFS), WiFi, and iBeacon signals were generated to gain insights into their spatial distribution across different floors. Two models were created: Model 1 integrated GFS, WiFi, iBeacon, and motion sensor data (accelerometer, gyroscope, and rotational), while Model 2 relied solely on GFS data. The results showed that Model 2, which utilized only the naturally occurring GFS, outperformed Model 1, achieving a Mean Absolute Error (MAE) of 36.72 meters compared to Model 1's 45.73 meters. This highlights the potential of geomagnetic data for indoor localization, as it is more stable and less influenced by environmental factors than WiFi and iBeacon signals.

However, while Model 2 demonstrated better overall performance, significant room remains for improvement. Some location predictions still deviate substantially from the actual ground truth. Future work will focus on enhancing the integration of sensor data (accelerometer, gyroscope, rotational) with GFS using advanced techniques like Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks to capture the temporal dynamics of movement more effectively. Transfer learning approaches will also be explored to adapt pre-trained models to new indoor environments, reducing the need for extensive retraining.

## IX. TEAM MEMBERS' CONTRIBUTIONS

### A. *Sathananthar Suresh*

- Essential Task 2: Visualize Geo-magnetic Heatmap
- Coding of visualization plot (visualize_f_matplotlib.py) using matplotlib and mpld3.
- Coding of the Bonus Task on Deep Learning-based location fingerprint model.
- Contributed to various project discussions.
- Compiled the full source code for the project into a single Jupyter Notebook.

### B. *Kee Ming Yuan*

- Contributed to typing the below sections of the report:
    - Introduction
    - Dataset including the pre-processing of WiFi
    - Essential Task 1: Visualization of Waypoints
    - Bonus Task 1: Build a Deep Learning-Based Location Fingerprinting Model
- Coded the maps for Essential Task 1: Visualization of Waypoints
- Helped in discussion and coding of Bonus Task 1

### C. *Robert George Susanto*

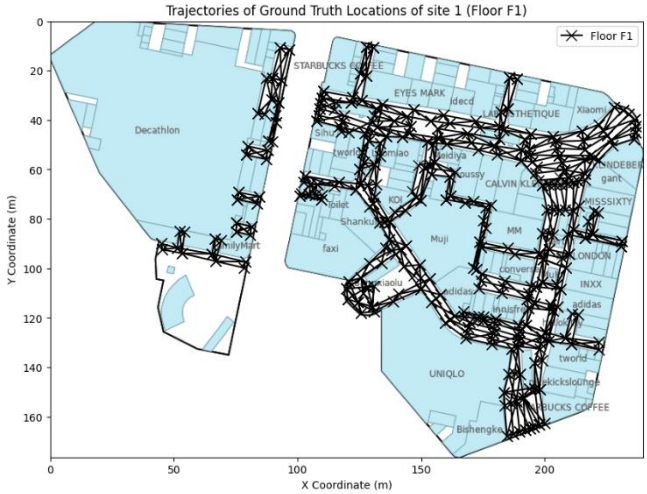
- Essential Task 4: iBeacon
- Coded the map for Essential Task 4: iBeacon
- Helped with Dataset Write up

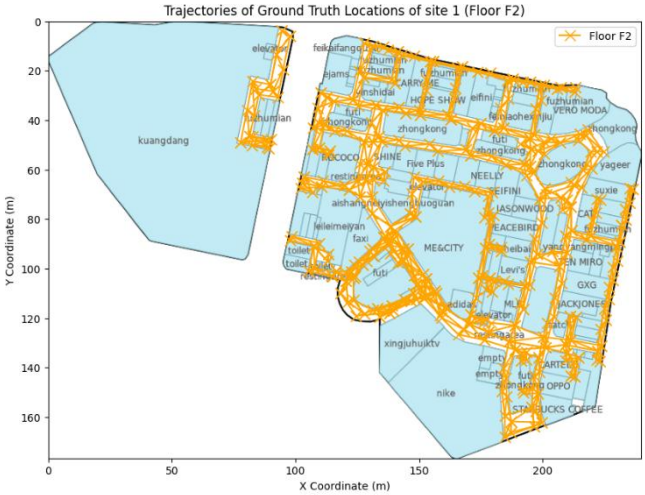
### D. *Sui Jinhong*

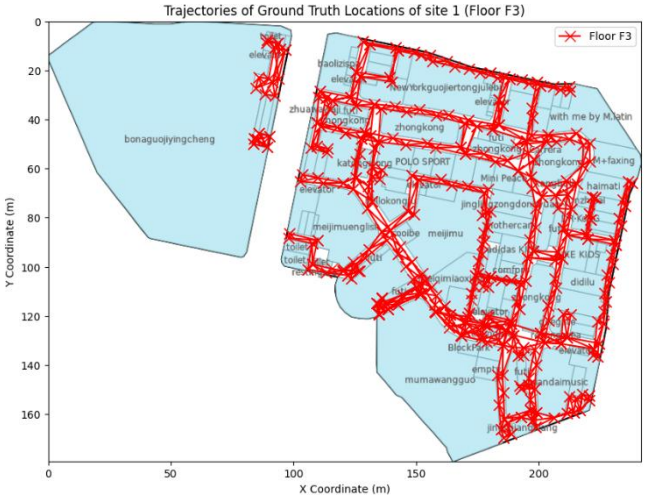
- Essential task 3: Coding, visualization and write up of wifi RSS
- Bonus task: Coding and write up of Bonus task on deep learning based location fingerprint model

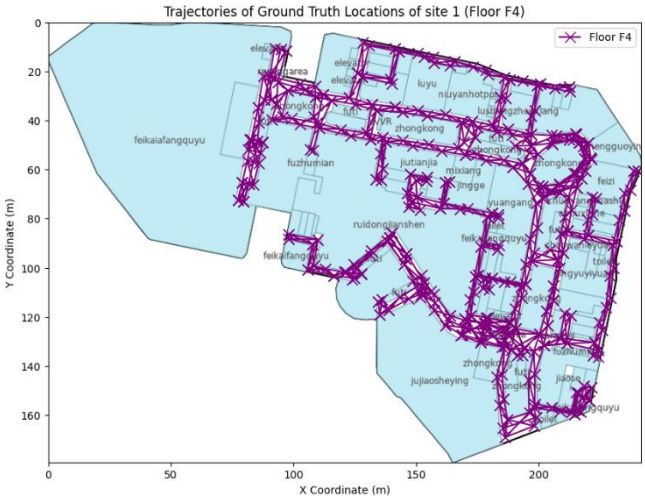
## X. APPENDIX

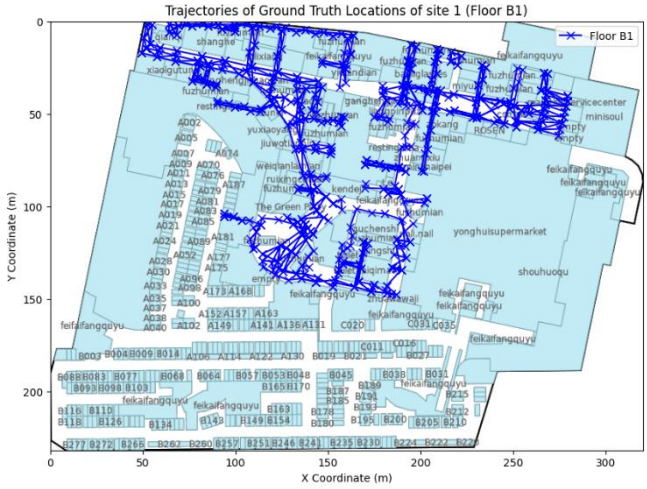
### SITE 1 WAYPOINTS

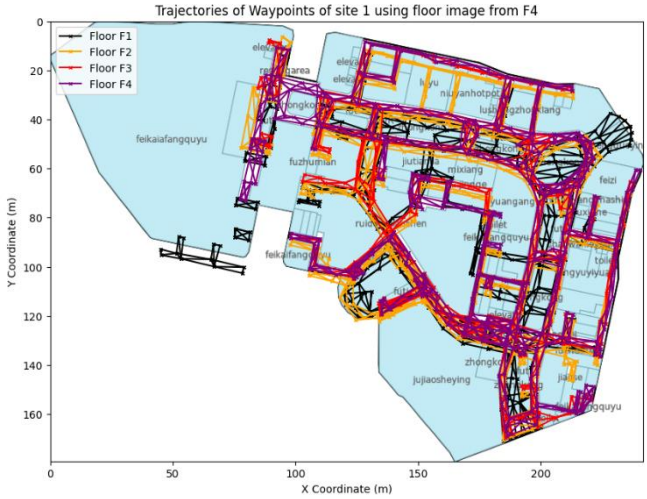


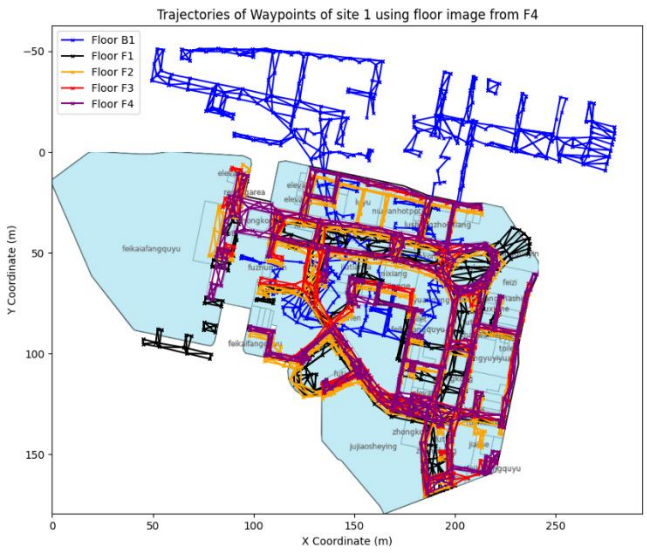



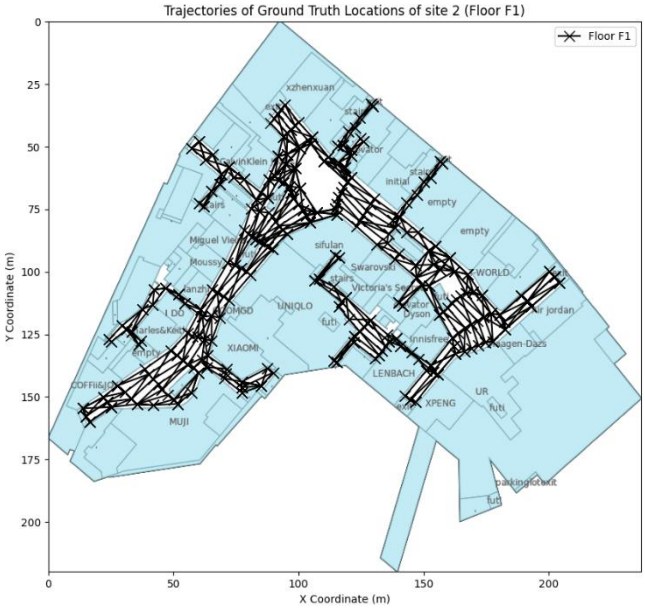



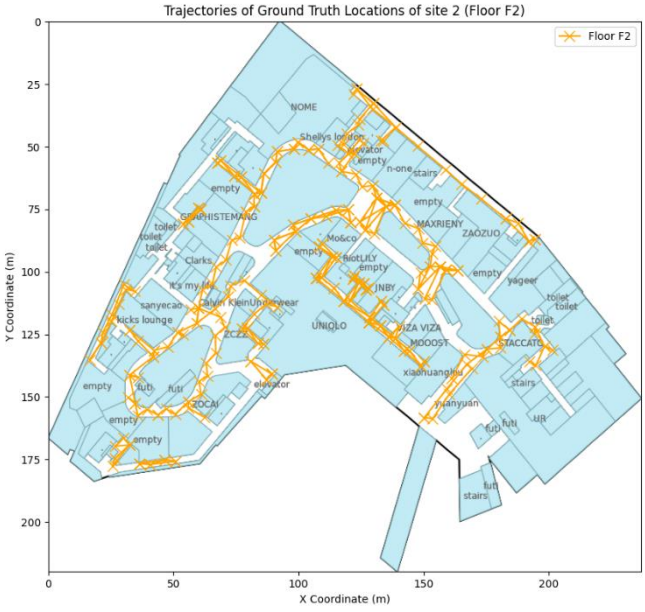


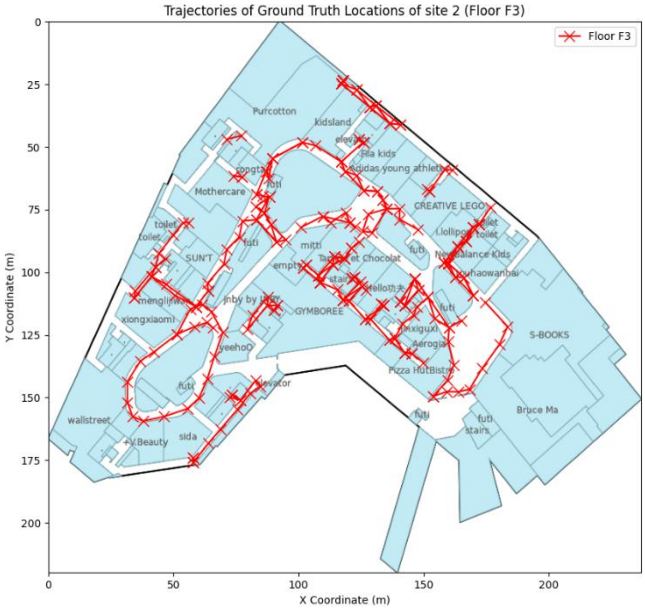



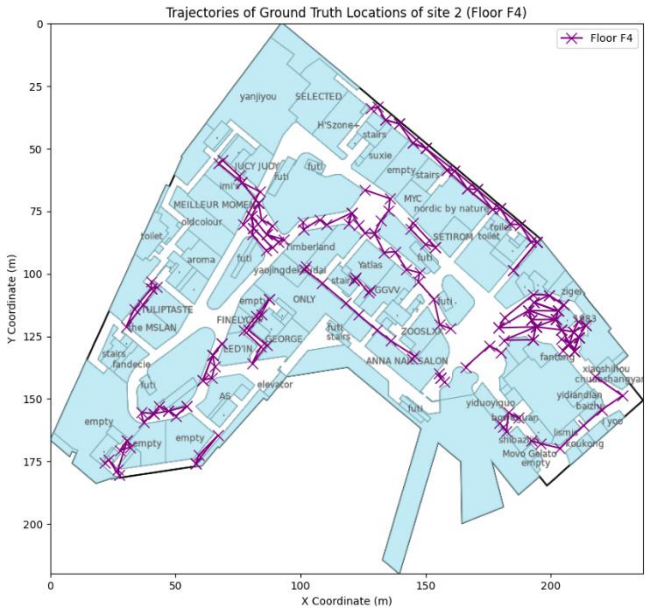



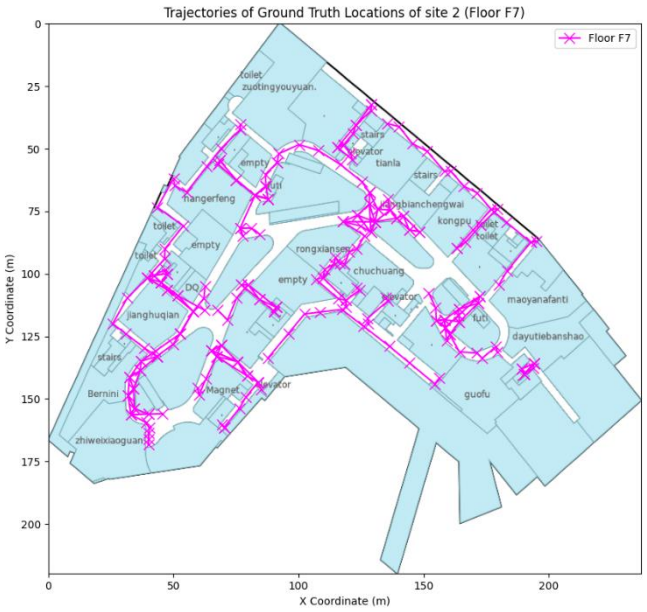

SITE 2 WAYPOINTS

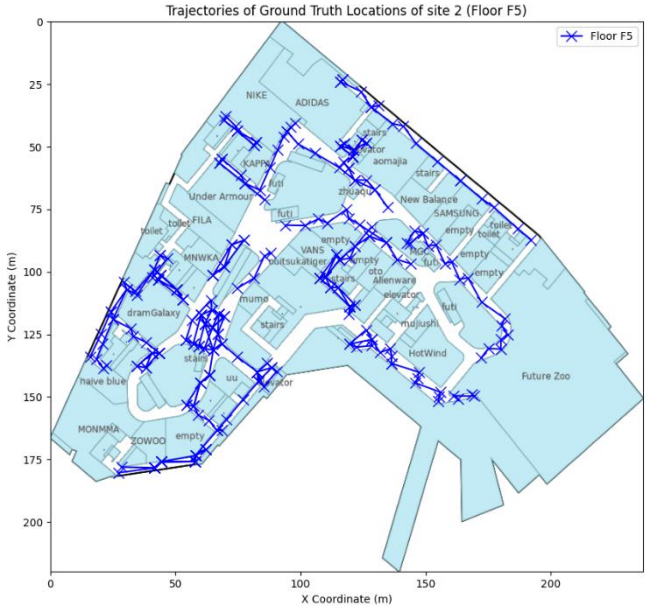


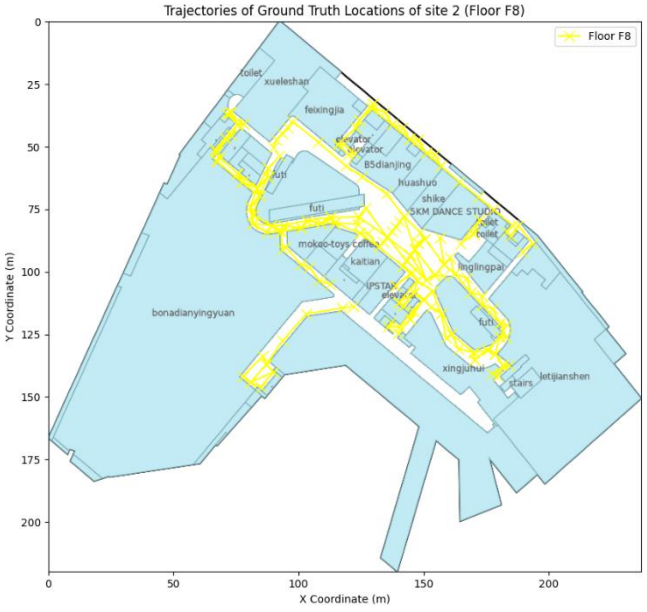



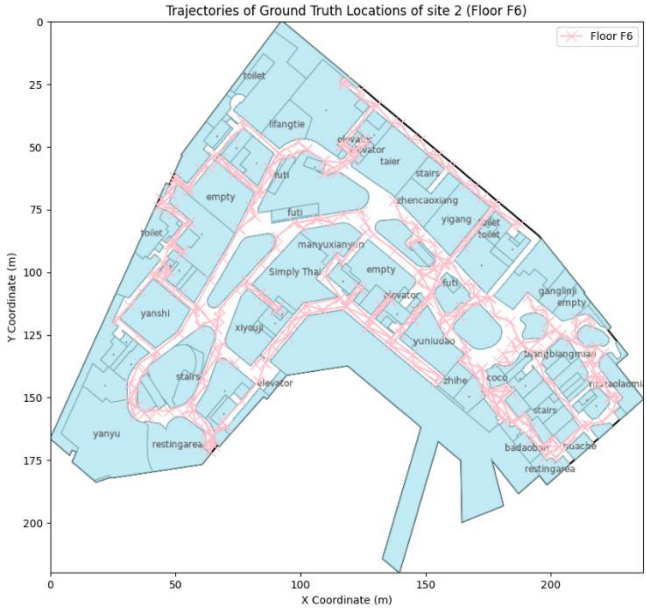



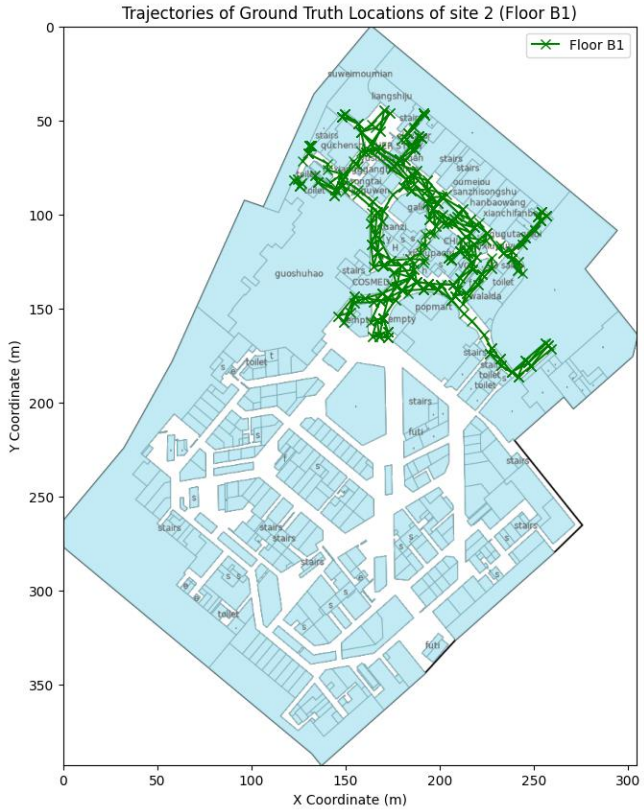

Trajectories of Ground Truth Locations of site 2 (Floor F4)
Floor F4
X Coordinate (m)
Y Coordinate (m)

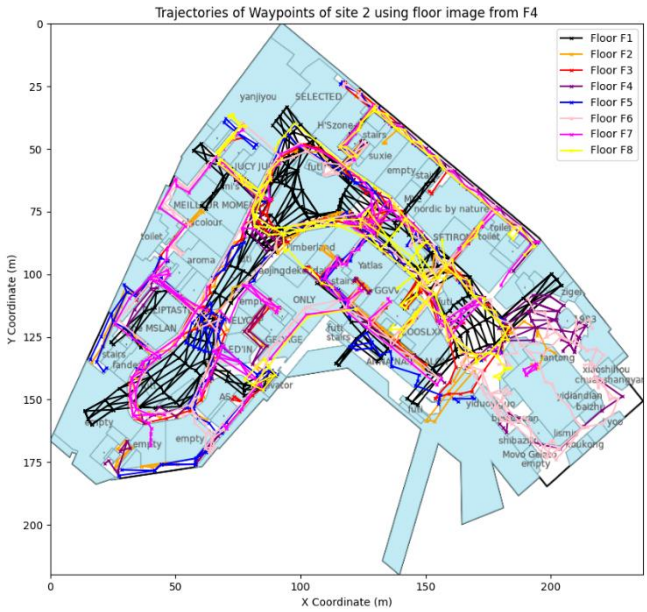

Trajectories of Ground Truth Locations of site 2 (Floor F7)
Floor F7
X Coordinate (m)
Y Coordinate (m)

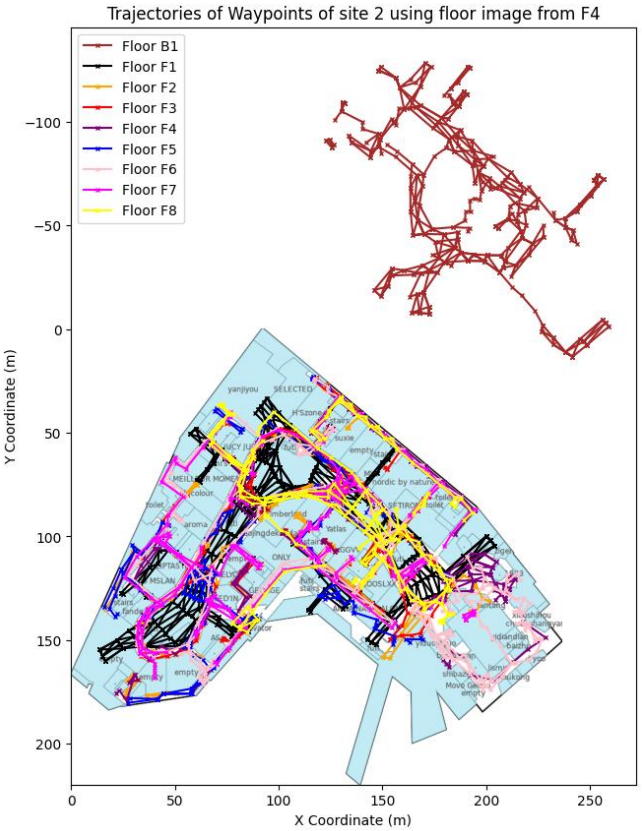

Trajectories of Ground Truth Locations of site 2 (Floor F5)
Floor F5
X Coordinate (m)
Y Coordinate (m)


Trajectories of Ground Truth Locations of site 2 (Floor F8)
Floor F8
X Coordinate (m)
Y Coordinate (m)


Trajectories of Ground Truth Locations of site 2 (Floor F6)
Floor F6
X Coordinate (m)
Y Coordinate (m)




Model specifications:

| Hyperparameter | Model 1: LocalizationNN | Model 2: DeeperMLP |
|---|---|---|
| Number of Layers | 4 (3 hidden + 1 output) | 3 (2 hidden + 1 output) |
| Input Size | X_train.shape[1] (based on training data) | X_train_scaled.shape[1] (based on training data) |
| Hidden Size (Layer 1) | 512 | 64 |
| Hidden Size (Layer 2) | 512 | 32 |
| Hidden Size (Layer 3) | 512 | N.A |
| Output Size | 2 (predicting x, y coordinates) | 2 (predicting x, y coordinates) |
| Activation Function | ReLU | ReLU |
| Dropout Probability | 0.3 | 0.4 |
| Batch Normalization | Yes (on all hidden layers) | Yes (on all hidden layers) |
| Loss Function | SmoothL1Loss | L1Loss |
| Optimizer | Adam | Adam |
| Learning Rate | 1.00E-03 | 1.00E-03 |
| Weight Decay | 1.00E-03 | N.A |
| Batch Size | 128 | 64 |
| Number of Epochs | 1000 | 1000 |
| Learning Rate Scheduler | ReduceLROnPlateau (factor: 0.1, patience: 10) | N.A |

## XI. References

[1] Google Colaboratory, "Indoor Location & Navigation Data Processing," Available: https://colab.research.google.com/drive/1z3EhIBNwuZbqffw36hizQ1I4iiVOl6d2, Accessed: Oct. 13, 2024.

[2] "IBeacon." *Wikipedia*, 14 Jan. 2023, en.wikipedia.org/wiki/IBeacon. [Accessed: Oct. 13, 2024].

[3] Apple Inc., "Getting Started with iBeacon," developer.apple.com/ibeacon. [Online]. Available: https://developer.apple.com/ibeacon. [Accessed: Oct. 13, 2024].