[Mobile and Ubiquitous Computing Mini Project] Step-by-Step Indoor Localization

Sangwi Kang
Seoul National University
Seoul, Republic of Korea
ksw292002@snu.ac.kr

Sangwi Kang. 2024. [Mobile and Ubiquitous Computing Mini Project] Step-by-Step Indoor Localization. In .
, 6 pages. https://doi.org/

1 INTRODUCTION

Localization methods are currently being used in various fields. Not only are they used at the government level in military, space, and aviation applications, but location-based services for commercial use are also actively being utilized. Specifically, Outdoor Localization is generally working smoothly with modern GNSS (Global Navigation Satellite System) such as GPS. However, there is currently no widely adopted solution for Indoor Localization. While Outdoor Localization can tolerate a certain degree of error without significantly impacting the service, even small errors in Indoor Localization can greatly degrade the quality of the service. The need for more precise accuracy makes it challenging to develop a generally applicable solution. Additionally, it is difficult to use methods that rely on receiving external signals like GNSS, as these signals often struggle to penetrate indoor environments effectively.

As large-scale indoor spaces that can accommodate diverse activities have become more commonplace, the need for Indoor Localization has increased significantly. Therefore, there is a pressing need for research on deployable and efficient Indoor Localization solutions.

In the current proposals for Indoor Localization, there are two main approaches that can be distinguished. One is the method of installing specific beacons and receiving their signals to determine the location. The other is the method of estimating the position based on the sensor data from the user's device.

The beacon-based approach can potentially achieve high accuracy, but there may be challenges in deployment. The user device sensor data-based approach, on the other hand, can enable device-free localization, but the accuracy may depend on the navigation algorithms used, potentially leading to errors.

In this mini project, we propose a method that estimates the position based on user device sensor data, while aiming to minimize the occurrence of errors. First, we will examine the research problem and identify the limitations of existing proposals. Then, we will introduce the approach taken in this project, describe the experimental procedure, analyze the experimental results, and discuss the limitations.

, ,

ACM ISBN
https://doi.org/

2 RESEARCH PROBLEM

For Indoor Localization, there are several approaches that can be used:

- Receiving specific indoor location signals to determine the position: This method involves the user's device receiving signals from the indoor environment to pinpoint the location.
- Dead reckoning using user device data: This approach utilizes sensor data from the user's device, such as accelerometers, gyroscopes, and magnetometers, to estimate the user's position through dead reckoning techniques.
- Fingerprinting based on wireless signals: This method relies on creating a database of wireless signal fingerprints (e.g., Wi-Fi, Bluetooth) at different locations, and then matching the user's current signal readings to the fingerprint database to determine their position.

These are some of the main techniques that have been proposed and explored for Indoor Localization. Each approach has its own advantages and challenges in terms of accuracy, deployment, and practicality, depending on the specific requirements and constraints of the indoor environment.

Among the various methods, the most realistically feasible solution is to deploy beacons and have the user's device receive the beacon signals to determine the user's location. This approach is practical because once the beacons are deployed and transmitting signals, the user's device can directly receive the signals and identify its position.

However, this method has a significant drawback - the need to deploy beacons across a wide area, which can incur substantial costs. Despite this cost issue, this beacon-based approach is considered the most realistically implementable solution when considering the practical constraints.

In other words, while the beacon-based method may not be the most optimal in terms of cost, it is the most realistic and feasible solution for Indoor Localization among the different approaches currently available.

"No Need to War-Drive: Unsupervised Indoor Localization."(H. Wang et al. MobiSys' 12. 2012) proposes a method called UnLoc, which primarily uses a dead reckoning approach. The dead reckoning method inherently introduces errors, which are initialized using LMs (Landmarks). The LMs are generated based on user device sensor data and contain location information. These LMs can be dynamically created as the user device moves. However, this solution cannot correct the errors that occur in the dead reckoning before encountering the LMs, and the real-time generation of LMs on the user device leads to an Energy Consumption issue.

The method of determining location based on wireless signal fingerprints involves measuring the wireless signals using the user device at a specific location and learning the pattern of these wireless signals to match them with a specific location. Commonly used wireless signals include Wi-Fi, Bluetooth, and magnetic fields, and recently, the use of Channel State Information (CSI) has gained popularity. Compared to using a single feature (such as signal strength), CSI can capture a much larger number of signal characteristics, enabling more accurate location estimation. However, this method is vulnerable to environmental changes, as variations in the surrounding environment can alter the wireless signal fingerprint at a specific location. This project also utilizes signal fingerprints of specific locations, and therefore, the team has attempted to address this issue as much as possible through a step-by-step approach.

3 SOLUTION APPROACH

3.1 Challenges

The currently proposed Indoor Localization methods have several challenges that need to be considered. This project aims to overcome these challenges and improve upon existing solutions in the following ways:

- This project does not use beacons. Deploying beacons requires significant costs, so to facilitate easy deployment of the solution, this project will use only user device sensor data for indoor localization.
- This project utilizes sensor data fingerprints of specific locations. However, due to the nature of fingerprint-based indoor localization, there is a high possibility of encountering duplicate fingerprints in a large-scale space. Therefore, this project proposes a step-by-step approach to narrow down the search space from a coarse Area Map to a more detailed Location Map, which helps address the issue of duplicate fingerprints.
- Collecting and learning the signal patterns at specific locations to create a signal pattern map is necessary. While this process incurs some costs, it can be done using only signals that can be collected from general user devices (such as Android phones) without the need for specialized signal generation equipment. The validity of using signal pattern maps for location identification has been verified in other research.
- Since this project requires collecting user device sensor data and creating signal pattern maps, it cannot create a solution that is suitable for all spaces and locations. Considering the scope of a mini-project, the data collection will be limited to a specific location, reducing the number of movement cases. After the experiment, the author will consider expanding the application of this solution to a wider range of spaces based on the results.

3.2 Solution Flow

Figure 1 illustrates the overall flow of this project. It shows a step-by-step approach to reduce the search space from the coarse Area Map to the more detailed Location Map. This process is intended to minimize the possibility of duplicate location fingerprints. Additionally, if the location fingerprint is directly searched within the entire space, the matching possibility may decrease due to the influence of the surrounding environment. Therefore, by first predicting the area and then finding the location within that specific area, the impact of the surrounding environment can be minimized.

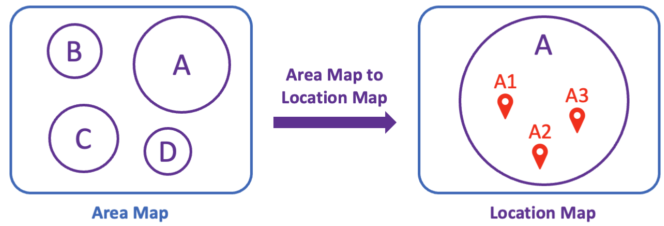

Figure 1: Area Map to Location Map

Both the area prediction and location identification require prior data collection and learning using the user device. The data to be collected and learned are as follows:

- Area Search Space Reduction: (Step Count, Moving Time, Turn Count, Last Rotation State) to Determine the Specific Area
- Location Fingerprint Matching: (Magnetic Field, Humidity) of the Specific Location

Among these, the Humidity data for Location Fingerprint Matching is omitted in this project due to the limitations of the user device used.

3.3 Prediction Step

The implementation of this project follows the three steps outlined below, after the prior collection and learning of the data mentioned for Area Search Space Reduction and Location Fingerprint Matching.

- Step 1: Moving. Collect the movement data from the same starting point to a specific Area.
- Step 2: Area Prediction. Predict the area code based on the measured data (Step Count, Moving Time, Turn Count, Last Rotation State).
- Step 3: Location Prediction. Predict the final location based on the Location Fingerprint (Magnetic Field [MagX, MagY, MagZ], AreaCode) recorded for the predicted Area Code.

4 EXPERIMENT

4.1 Data Recording

For the progress of this project, the collection of moving data for area prediction and magnetic data for location prediction is necessary. An Android Phone (Samsung Galaxy S22 / Android 14) was used to collect the data, utilizing the Step Detector, Gyroscope, and Magnetic Field Sensor on the device.

To collect the data, a Data Recording Tool Application was developed for the device mentioned above. Figure 2 shows the developed Data Recording Tool. This app collects Step Count and Gyroscope

Figure 2: Step-by-Step Indoor Localization Data Recording Tool - Android

data through the Step Count feature, and Magnetic data through the Magnetic feature.

To record the data for a specific area, the user can start the app, press the Step Count button, and move from the same starting point to the target area. The app will record the (Step Count, Moving Time, Turn Count, Last Rotation State) for the journey to that area.

To collect the location fingerprint, the user can stand at a specific location, press the Magnetic button, and the app will record the Magnetic Field Data (MagX, MagY, MagZ) for that location over a certain period. Since the experimenter knows which area the location fingerprint is being collected in, the LocationCode is manually added to the collected Magnetic data.

The collected data is saved as a .csv file in the internal storage of the Android Phone.

The Moving Data for specific Areas was measured through 25 repeated trials. The Fingerprint for specific Locations was measured 150-200 times over a certain time period. Both the Moving Data and Location Fingerprint were measured in the Hand-Held state, as shown in Figure 3.

Figure 3: Hand-Held Data Recording

4.2 Area and Location

The map for the experiment is shown in Figure 4. Figure 4 depicts the map of the space where the experiment was conducted. The target space is the 5th floor of the Department of Computer Science and Engineering graduate student shared research lab in Building 301 at SNU. The target Areas are A, B, C, and D, totaling 4 areas. The target Locations are A1 D2, totaling 8 locations. The descriptions of each Area and Location are as follows:

- A1 : 551-2
- A2 : 554-1
- B1 : Lounge
- B2 : Men's Toilet 1
- C1 : Pantry
- C2 : Maldives Meeting Room
- D1 : 512-2 Door
- D2 : Men's Toilet 2

The criteria for dividing the Areas and Locations was based on the expected similarity of the Location Fingerprints. Locations that were expected to have similar Location Fingerprints were separated into different Locations. Areas that were adjacent spaces were considered as part of the same Area.

The data measured based on the defined Areas and Locations is shown in Figure 5. Figure 5 presents the Moving Data for Area identification and the Magnetic Data for Location identification.

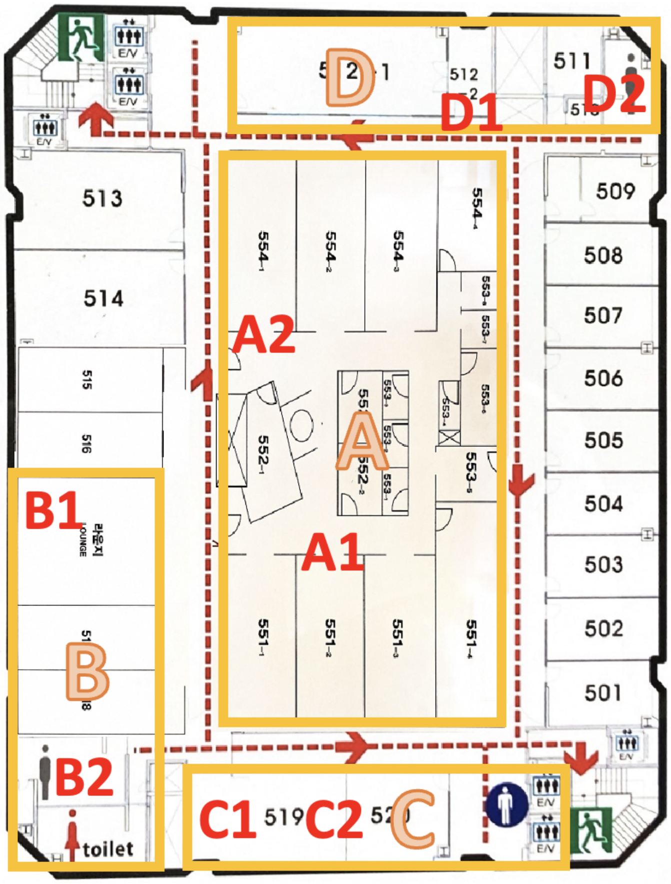

Figure 4: Experiment Target Map

Timestamp	StepCount	MovingTime	TurnCount	LastRotation	Area
1.71767E+12	42	25180	1	1.0070106	A
1.71767E+12	33	18739	1	0.870406	A
1.71767E+12	33	19303	1	0.909654	A
1.71767E+12	43	25395	1	1.159345	A
1.71767E+12	40	21021	1	0.8035926	A

Timestamp	MagX	MagY	MagZ	AreaCode	LocationCode
1717672487027	18.075	-30.562502	-20.5125	A	A1
1717672488017	17.90625	-31.650002	-21.375	A	A1
1717672489023	18.018751	-32.081253	-21.525002	A	A1
1717672490012	17.925001	-32.4375	-21.46875	A	A1
1717672491021	17.85	-32.68125	-21.4125	A	A1

Figure 5: Moving Data(Up) and Magnetic Data(Down)

4.3 Training Model

The Training Model used is the Random Forest model, which can effectively classify non-linear data. The Moving Data and Magnetic

	Precision	Recall	F1-Score
A	0.67	1.00	0.80
B	1.00	1.00	1.00
C	1.00	1.00	1.00
D	1.00	0.82	0.90
Accuracy			0.93

Figure 6: Area Prediction Result

	Precision	Recall	F1-Score
A1	1.00	0.94	0.97
A2	0.94	1.00	0.97
B1	0.87	0.91	0.89
B2	0.92	0.89	0.90
C1	0.95	1.00	0.98
C2	1.00	0.95	0.97
D1	0.86	0.88	0.87
D2	0.87	0.86	0.86
Accuracy			0.93

Figure 7: Location Prediction Result

Data were input into the model with the following Features and Labels for training:

Area Prediction :

- Feature : (StepCount, MovingTime, TurnCount, LastRotation)
- Label : Area

Location Prediction :

- Feature : (MagX, MagY, MagZ, AreaCode)
- Label : LocationCode

5 EVALUATION

Based on the measured data (Moving Data, Magnetic Data), the results of evaluating the trained model are as follows:

Figure 6 shows the results of predicting the Areas (A, B, C, D) after training the Moving Data. The overall accuracy achieved is 0.93, indicating a reasonable level of accuracy. However, the precision in predicting Area A is 0.63, which suggests the need for further improvement in that area.

Figure 7 shows the results of predicting the Locations (A1 D2) after training the Magnetic Data. The overall accuracy achieved is 0.93, which is at the same level as the Area Prediction. However, the precision and recall for predicting the Locations in Area D (D1, D2) tend to be lower.

Both the Area Prediction and Location Prediction achieved an accuracy of 0.93, which is a usable level of accuracy. However, considering the limited size of the overall dataset and the restricted movement, additional data collection for the areas and locations is necessary.

6 LIMITATION

The goal of this project was to propose an non-error indoor localization method using a device-free approach without the need for beacons. However, the experimental environment was highly limited, and the following limitations exist, which require further improvements in the future:

6.1 Humidity Sensor Issue

When conceptualizing this project, the intended sensors included a humidity sensor, as the humidity data was expected to play an important role in differentiating locations. The use of Wi-Fi or Bluetooth signals as the primary fingerprints was also considered, but these signals are from devices that many people carry, and as people enter and exit specific locations, the signals are repeatedly added and removed, making them highly susceptible to environmental influences.

In contrast, humidity data can be affected by weather conditions, but in a controlled indoor environment like a research lab with temperature and humidity regulation, the impact is expected to be minimal. Locations with significant humidity differences, such as toilets and pantries, were anticipated to provide useful location-based humidity data. Therefore, humidity data was expected to be suitable for Location Fingerprinting.

However, the device used in this project, the Samsung Galaxy S22, does not have a built-in humidity sensor, making it impossible to measure the humidity data. Further investigation revealed that while past Android devices had included humidity sensors, most recent Android phones no longer have this feature. Therefore, to measure the humidity data, a separate humidity sensor device would be needed.

6.2 Starting Point and Moving Path Issue

The current data measurement approach involves moving from a specific starting point to a specific destination point while collecting data. In this method, the starting point must always be the same. If the starting point is not consistent, the proposed method in this project cannot accurately predict the Area.

Of course, the general movement patterns (within a research lab) were considered as much as possible when determining the starting point and measuring the Moving Data. However, not everyone follows a typical movement pattern. People can make sudden, unexpected movements or change their destination abruptly. Moreover, the current scenario assumes the ability to reach a specific Area after entering the building, but if the starting point is different, the Moving Data (Step Count, Moving Time, Turn Count, Last Rotation) will be completely different, making accurate prediction impossible.

This is the biggest issue with the current project and a crucial aspect that requires further improvement.

6.3 Insufficient dataset

The lack of dataset can also be considered a limitation. For this project, an Android app was developed, and an Android phone was used to measure data while walking around in a hand-held state. Collecting the Location Fingerprint data, such as Magnetic Data, only required holding the Android phone at a specific location, and the data was collected automatically, without incurring significant costs.

The measurement of Moving Data for Area Prediction required an enormous amount of time. As the area was divided, data had to be measured for each divided area, and repetitive data measurement was necessary for model training. In this project, data was measured 25 times. The problem was that this method assumed the same starting point, so it was necessary to return to the same starting point after each measurement. This process was very time-consuming and physically demanding.

The method used in this project clearly had limitations in collecting sufficient data. It is necessary to come up with other measurement ideas, such as automating data collection, to get rid of this problem.

6.4 Unclear Distinctions of Areas and Moving Path

Another limitation is that the criteria for distinguishing Area, Location, and Movement Paths are not clear. In this project, Area, Location, and Movement Paths were distinguished based on human intuition. Of course, since the project considered the most general movements and human activities as much as possible, the proposed method in this project is likely to work with a certain level of reliability. However, it is estimated that the method worked to some extent because the areas, locations, and movement paths were divided into places and paths that could show as different characteristics as possible.

In order for the experiment to work in general situations rather than limited situations, more extensive data on Area, Location, and Movement Paths needs to be collected. However, it is not possible to manually divide the locations and set the movement paths based on human intuition every time. It is absolutely necessary to establish reasonable criteria and automatically classify them.

7 CONCLUSION

In this project, we identified the reasons and limitations of the lack of a general solution for indoor localization, and proposed a new method to implement indoor localization in a novel way. The most promising beacon-based approach proposed previously had the problem of high deployment costs, while the methods based on user device sensor data had issues with estimation algorithm errors and high sensitivity to the surrounding environment.

The Step-by-Step Indoor Localization method proposed in this project eliminates the need for beacons, thus avoiding the deployment costs associated with them. It also avoids the errors that occurred in the widely used estimation algorithms. By progressively reducing the search space towards the final Location, the influence of the surrounding environment is also minimized. When applying the Step-by-Step Indoor Localization to 4 Areas and 8 Locations on a single floor, the project achieved an accuracy of 0.93 for both Area Prediction and Location Prediction, reaching a usable level of accuracy.

However, the Step-by-Step Indoor Localization approach has some limitations. Firstly, the division of Areas, Locations, and trajectories was based on human intuition, which may not be scalable or generalizable. Secondly, the collected data was only valid in very restricted situations, and thirdly, the data collection process was extremely time-consuming.

Therefore, to address these issues, further improvements are necessary. This includes:

- Applying additional sensor data to enhance the approach.
- Establishing clear and objective criteria for the division of Areas, Locations, and trajectories, rather than relying on human intuition.
- Automating the data collection process to make it more efficient and scalable.

These improvements are essential to overcome the limitations of the current Step-by-Step Indoor Localization method and develop a more robust and generalizable solution for indoor localization.

8 MATERIALS

The code for the Data Recording Tool, code for the Data Training Implementation, dataset, brief presentation, and other materials used in this project can be accessed at https://github.com/snudev-swkang/sbs-indoor-localization.git.