

A Simplified Multi-Floor Classification-Based Indoor Positioning System Study

No Author Given

No Institute Given

Abstract. Indoor Positioning Systems (IPS) aim to supplement or replace Global Navigation Satellite Systems (GNSS) for indoor localization, yet limited research has addressed IPS optimization in multi-floor environments. This study investigates the overall performance of machine learning-based fingerprinting models under varying environmental configurations, with a focus on two key factors: grid size for fingerprint data collection and the impact of low-relevance Basic Service Set Identifiers (BSSIDs). Two new evaluation metrics, which are Average Grid from Target (AGT) and Average Distance from Target (ADT), are introduced to provide standardized measures of positioning accuracy across different grid layouts. To reduce computational load and training time, a BSSID filtering approach is applied, excluding low-intensity signals likely to be irrelevant to localization. Experiments were conducted in the hallways of a university learning center across two floors, with an initial $1 \times 1\text{m}$ grid size. The dataset comprised 12,640 RSSI samples collected from 378 filtered access points over approximately 500 grid cells. Larger grid sizes were synthesized via interpolation to evaluate performance trade-offs. Results show that while accuracy generally increases with grid size, a $7 \times 7\text{m}$ configuration achieves the optimal balance, yielding 80% accuracy alongside the lowest AGT and ADT values. Random Forest and XGBoost consistently outperformed other models, indicating their suitability for multi-floor IPS tasks in similar environments.

Keywords: Indoor Positioning System (IPS) · Wi-Fi Signal Processing · Machine Learning · Multi-Floor Positioning

1 Introduction

Indoor Positioning Systems (IPS) are designed to help users determine their location and navigate effectively within buildings or enclosed areas. Traditional positioning and navigation methods, such as the Global Navigation Satellite System (GNSS), rely on signals from satellites and ground control stations to calculate positions using trilateration. The most well-known GNSS is the Global Positioning System (GPS). However, GPS performance degrades significantly indoors because most buildings are constructed from dense materials like concrete and metal, which block or weaken the low-power satellite signals. This results in scattering, shadowing, blind spots, and signal attenuation, all of which reduce positioning accuracy [11]. To overcome these limitations, various IPS

methods have been developed to provide accurate navigation in environments where GNSS is unreliable or unavailable.

Many methods have been developed for Indoor Positioning Systems (IPS), with one of the most well-known being trilateration. This technique measures the signal strength from transmitters, such as Bluetooth Low Energy (BLE), and uses the Received Signal Strength Indicator (RSSI) as a proxy for distance [3]. The process typically involves an offline phase to create an RSSI “radio map” (fingerprinting), followed by an online phase to estimate the user’s position based on those measurements. However, a major drawback is that RSSI measurements are susceptible to environmental noise and require the precise placement of access points, which ultimately limits their accuracy [14].

Machine learning approaches have proven effective for IPS performance improvement [6]. Classification algorithms, e.g., Support Vector Machine (SVM), k-Nearest Neighbor (kNN), Random Forests, and Neural Networks address RSSI limitations by treating positioning as a classification problem, providing reliable area-based location estimates.

Previous research [7] implemented an IPS using a large grid size ($16.75 \times 15\text{m}$) to reduce the number of classification labels, making the problem more manageable within time constraints. While this approach provided a functional implementation, it left several questions unanswered regarding data point influence on IPS performance, the feasibility of implementing reliable IPS with limited Basic Service Set Identifier (BSSID) features, and the minimum viable grid size for maintaining positioning accuracy. Building on this foundation, this paper further expands classification-based IPS by refining the approach to ground truth reliability in experimental settings. The main contributions of this paper are:

1. Introduction of new evaluation metrics: Average Grid from Target (AGT) and Average Distance from Target (ADT).
2. Investigation of trade-offs between grid size and precision, analyzing implementation advantages and limitations.
3. Presentation of deeper insights into IPS design considerations, offering practical improvements for indoor positioning accuracy.
4. Examination of feature filtering effects on model complexity.

The remainder of this paper is organized as follows: Section 2 presents a literature review to provide an overview of the research landscape; Section 3 describes the research methodology; Section 4 reports the experimental results; Section 5 discusses the technical findings; and Section 6 concludes the study with future directions.

2 Literature Review

With the aforementioned limitations of GNSS and GPS in indoor environments, research on IPS has gained significant momentum in recent years. Consequently, IPS has been extensively studied and developed using various signal processing

techniques [12]. One prominent approach is Wi-Fi-based RSSI fingerprinting, enhanced by machine learning (ML) and deep learning methods.

Traditional ML algorithms have been employed to classify fingerprints and estimate the current position from a set of RSSI values. Commonly used algorithms include k-Nearest Neighbor (kNN) [7, 15, 16], Random Forest [7, 9, 16], Support Vector Machine (SVM) [4, 7, 15, 16], and Multi-Layer Perceptron (MLP) [7, 15]. These models offer straightforward implementation and produce interpretable results. Their relatively simple architectures and modest computational requirements make them particularly suitable for IPS applications, as they can be deployed without specialized hardware. The accuracy of an IPS largely depends on the quality of the collected data, which can be influenced by factors such as diverse RSSI readings at different times [2], the integration of human activity recognition [1], and the use of coordinated data collection applications [10].

Deep learning approaches, such as Graph Neural Networks (GNNs) [15] and Convolutional Neural Networks (CNNs) [1], have demonstrated superior performance over traditional ML algorithms when sufficient data points are available. However, despite these advancements, key environmental variables, such as the resolution of fingerprint grids and the spatial distribution of reference points, have not been comprehensively investigated. In this study, we emphasize the critical role of these factors in IPS performance and systematically analyze the most effective configurations for multi-floor environments. By addressing these overlooked variables, our work aims to bridge existing research gaps and enhance the accuracy, reliability, and overall performance of IPS through optimal design and configuration.

3 Research Methodology

This paper adopts a systematic and empirical research methodology with the aim of identifying environmental factors that influence the IPS. This methodology is illustrated in Figure 1, consisting of four parts: (1) Environmental factor identification, (2) Data extraction, (3) Model training and inference, and (4) Result analysis. Each part will be explained in detail in the following sections.

3.1 Identification of Environment Factors

Environmental factors refer to any physical or logical characteristics of the area used by IPS developers to generate an RSSI-based radio map. In a typical offline map generation phase, the target area is marked and divided into grids that correspond to its digital floor plan. Each grid is assigned a unique reference number. Data collection is carried out using devices [8, 13] capable of measuring RSSI values from various reference points, such as BLE beacons or Wi-Fi access points, at each grid location. The resulting set of RSSI values serves as the feature vector representing each grid (or sub-grid) in the radio map. Sometimes, data synthesis is also used to add more data points from the collected fingerprints [16].

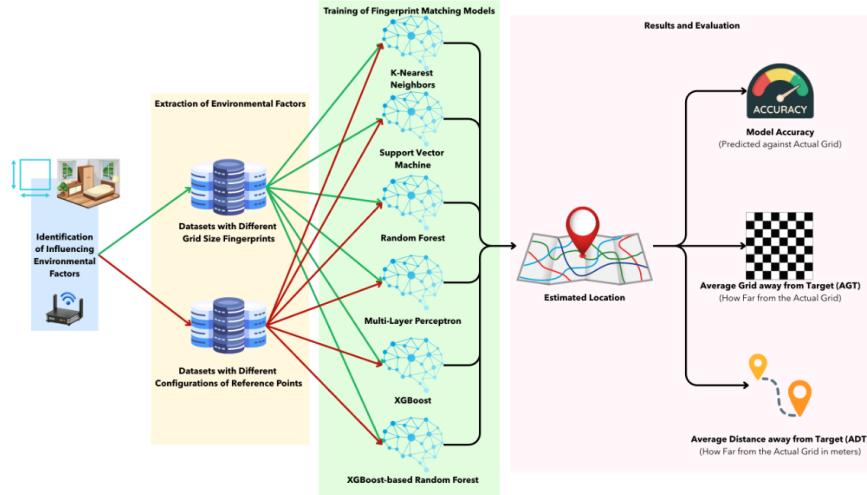


Fig. 1. An Overview of the Research Methodology.

From the analysis of the offline map generation process, two key factors emerge as potentially significant to IPS performance. First, grid size directly impacts the precision of position estimation: larger grids make it easier to classify a position into the correct grid, but smaller grids are required for higher positioning accuracy. However, reducing grid size substantially increases the labor required for data collection, rendering map generation impractical for large-scale environments. Second, RSSI collection is not strictly limited to reference points (e.g., access points) within the target area; signals from access points in neighboring buildings may also be recorded, introducing low-relevance data as shown in Figure 2. The x-axis represents the BSSID of different access points, with darker colors indicating lower received signal strength in decibel-meters (dBm). This study focuses on these two factors as the scope of investigation.

3.2 Extraction of Environmental Factors and Datasets

The two factors identified in the previous section can be tuned, and each is expected to have an optimal balance between practicality and performance. In this section, we extract the quantitative representation of each factor and prepare the dataset for model training and inference accordingly.

For the grid size factor, we quantify it using distance measurements. Traditionally, positioning evaluations rely on precise global coordinates (e.g., latitude/longitude) as benchmarks. However, such measurements are impractical to obtain indoors without professional surveying equipment. To address this, we introduce two new metrics based on the Euclidean distance concept: Average Grid from Target (AGT) and Average Distance from Target (ADT). The AGT

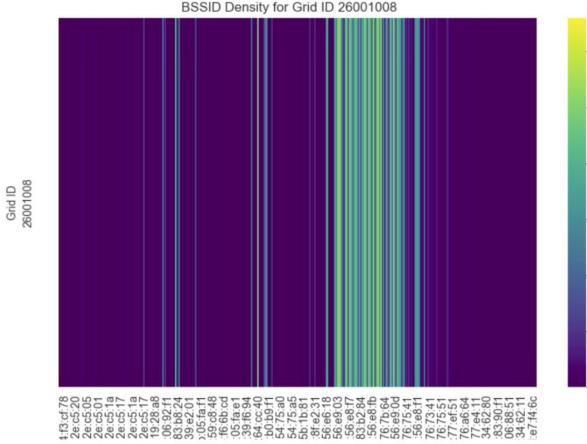


Fig. 2. A Heatmap of RSSI values collected from different access points.

can be computed using the following equation:

$$AGT = \sqrt{(x_{target} - x_{estimate})^2 + (y_{target} - y_{estimate})^2} \quad (1)$$

where x_{target} and y_{target} are the indices of the target grid on the horizontal and vertical axes, respectively, and $x_{estimate}$ and $y_{estimate}$ are the indices of the estimated grid predicted by the IPS. For example, $x_{target} = 3$ and $y_{target} = 5$ indicate that the target grid is the third grid from the left and the fifth grid from the top of the map. By treating grid indices as distance units, this metric naturally adapts to different grid sizes, providing both a straightforward means to visualize positioning performance and a standardized measure for cross-comparison across varying grid sizes.

On the other hand, the ADT is a related metric that quantifies the average Euclidean distance, measured in actual physical units such as meters, between the estimated position and the ground-truth target position. It can be computed by the following equation:

$$ADT = \sqrt{[w_g \times (x_{target} - x_{estimate})]^2 + [h_g \times (y_{target} - y_{estimate})]^2} \quad (2)$$

where w_g and h_g denote the global grid width and height in meters, respectively. Unlike AGT, which uses grid indices as units, ADT reflects the absolute positioning error in a real physical space, making it more directly interpretable for practical deployment. Notably, larger grid sizes inherently lead to greater distances between the target and the estimated positions, even when the grid index difference remains the same.

On the other hand, quantifying the low-relevance data factor is more straightforward. The performance of an IPS is inversely proportional to the amount

of low-relevance data, meaning that the more low-relevance data present, the less accurate the IPS becomes. To identify such data points, we apply a simple threshold-based approach. From the analysis of the collected RSSI heatmap in Figure 2, it is evident that the majority of low-relevance BSSIDs has its RSSI values fall within the range of 0 to 5 decibel-meters. Based on this observation, we propose using a threshold of 5 decibel-meters to filter out data points that are likely to negatively impact IPS performance.

With the metrics above, we can quantify the effect of different grid sizes in a tangible way. Ideally, data points for various grid sizes should be collected to study the influence of this factor. However, collecting new datasets for each grid size is highly impractical. In typical IPS experiments, floor taping is used to physically mark grid boundaries, and repeatedly re-taping the same area would be extremely time-consuming and prone to human measurement errors. To address this limitation, we propose an interpolation-based approach, where larger grids are generated by aggregating smaller ones. For example, four 1×1 m grids can be combined into a single 2×2 m grid, as illustrated in Figure 3. In each aggregated grid, RSSI values from five sampling points, i.e., top-left, top-right, center, bottom-left, and bottom-right, are interpolated using a standard averaging method.

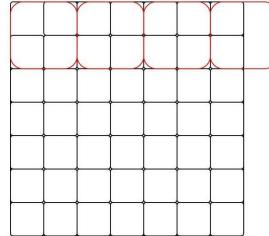


Fig. 3. Visualizing the process of grid aggregation by interpolating RSSI values.

A major drawback of the grid interpolation process is the reduction in the number of data points. However, from multiple rounds of data collection in this study, we obtained 9,019 and 3,621 data points from the first- and second-floor hallways, respectively. We believe this is a sufficient number of data points to maintain a high-quality dataset. The difference in the number of data points is due to variations in the floor plan and room partitioning of the test area. Table 1 summarizes the metadata and statistics of the dataset collected in this study.

Moreover, the data points in the datasets were analyzed to filter out low-relevance features. These features correspond to access points (referred to as BSSIDs), which are a set of RSSI values. Out of the 1,799 BSSIDs initially detected by the data collector device, 1,421 were filtered out as being low-relevance. At this stage, the datasets are ready for the model training process.

Table 1. Metadata of the number of grids and data points from different grid sizes.

| Floor | | Grid Sizes | | | | | | | |
|-------|--------------|------------|-------|-------|-------|-------|---------|---------|---------|
| | | 1 × 1 | 3 × 3 | 5 × 5 | 7 × 7 | 9 × 9 | 11 × 11 | 13 × 13 | 15 × 15 |
| 1st | # Grid | 309 | 55 | 28 | 17 | 17 | 10 | 10 | 5 |
| | # Data Point | 8,086 | 1,439 | 733 | 445 | 399 | 262 | 235 | 130 |
| 2nd | # Grid | 174 | 41 | 19 | 12 | 11 | 6 | 8 | 4 |
| | # Data Point | 4,554 | 1,073 | 497 | 314 | 288 | 157 | 209 | 105 |

3.3 Model Training

To study the impact of the identified environmental factors, this paper trains various machine learning models to estimate the current position in terms of the predefined grid layout.

The selection of algorithms was based on an analysis of the dataset. We found that each data point contains four main features—the current floor, grid size, current grid index, and a set of corresponding RSSI values—and that most of them are well-structured, sparse, numeric, and ordinal. Therefore, the chosen algorithms needed to be capable of handling and analyzing this type of data.

This paper uses four traditional machine learning algorithms (kNN, Random Forest, MLP, and SVM). Additionally, two tree-based ensemble models were selected: XGBoost [5] and a hybrid model combining XGBoost + Random Forest. Each model was trained through a standard process, with hyperparameter tuning performed to achieve optimal performance.

4 Experiments and Results

In this section, we conducted experiments to evaluate the performance of various ML algorithms and models on the IPS task. These models were trained as described in Section 3.3. First, we evaluated the overall performance of the models using key metrics such as accuracy, AGT, and ADT. Afterward, we analyzed the changes in accuracy and AGT that resulted from applying BSSID filtering.

4.1 Accuracy, AGT, and ADT Results

Every ML model is trained with the same controlled dataset and workstation. Then, we calculate accuracy, AGT, and ADT from the predicted test results compared to the ground truth in the dataset. We plotted graphs to visualize the result comparison between different models and grid sizes based on the accuracy, AGT, and ADT, as shown in Figures 4, 5, and 6.

Figure 4 presents the position estimation accuracy of each model across different grid sizes. The results indicate that larger grid sizes generally yield higher accuracy. However, we identified an optimal configuration where the smallest grid size that still achieves high accuracy (approximately 80%) is 7 × 7m. Additionally, Random Forest (RF) and XGBoost consistently deliver the best performance for position estimation across all grid sizes.

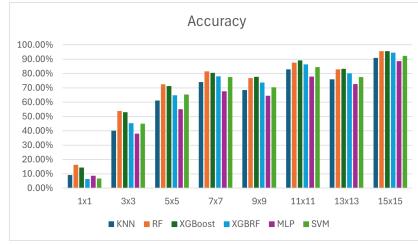


Fig. 4. Model Accuracy on Different grid sizes and models

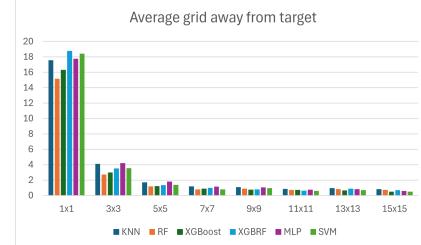


Fig. 5. Model AGT on Different grid sizes and models

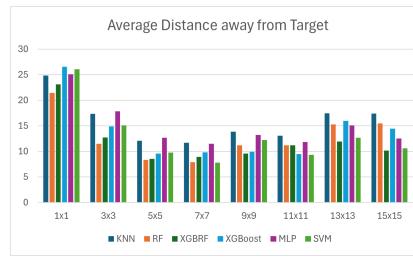


Fig. 6. Model ADT on Different grid sizes and models

For AGT and ADT in Figures 5 and 6, the results are similar to those for accuracy. Figure 5 shows that larger grids have lower AGT, meaning it is easier to classify the correct grid during position estimation. The AGT value begins to stabilize once the grid size reaches 7×7 m, suggesting that this is the smallest grid size capable of maintaining minimal errors. On the other hand, Figure 6 indicates that a 7×7 m grid size minimizes error while remaining small enough for precision-dependent applications in this specific setting. The ADT results confirm that, even when measured in physical units, the 7×7 m grid size remains the optimal point.

4.2 Filtering Impact on Accuracy and AGT measurement

As mentioned in the last paragraph of Section 3.2, BSSID filtering is proposed to exclude low-relevance data from the dataset. This approach can reduce computational resource requirements, particularly in scenarios with a large feature space (i.e., a high number of BSSIDs within the area network). To evaluate whether this filtering process can improve IPS performance, we conducted an experiment comparing the best-performing model—the Random Forest algorithm—trained on datasets generated with and without filtering.

Figures 7 and 8 depict the accuracy differences without and with the filtering process. The results show that applying the filtering improves the overall accuracy of the random forest model for position estimation from approximately 0.14 to 0.16 (about a 2% improvement). Furthermore, as shown in Figures 9

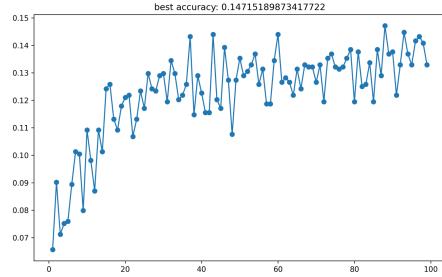


Fig. 7. Accuracy of random forest **without** BSSID filtering applied (1×1)

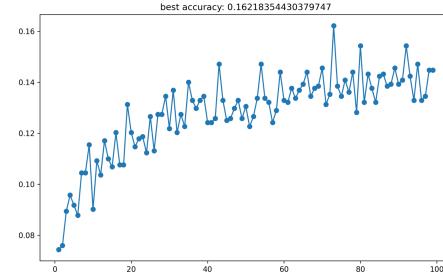


Fig. 8. Accuracy of random forest when we are **applying** BSSID filtering (1×1)

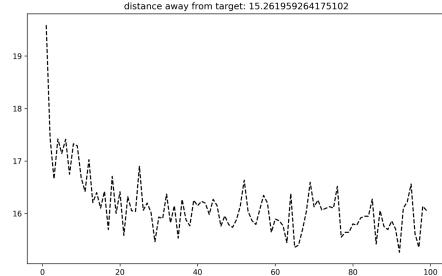


Fig. 9. AGT of random forest **without** BSSID filtering applied (1×1)

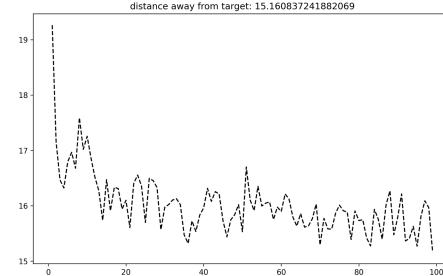


Fig. 10. AGT of random forest when we are **applying** BSSID filtering (1×1)

and 10, the filtering process does not negatively affect the AGT performance of the IPS model. In fact, AGT decreases from 15.28 to 15.18 meters, indicating a reduction in positioning error compared to the non-filtered model. It should be noted that the relatively low accuracy and AGT scores compared to Section 4.1 are due to the evaluation being conducted with only the first 100 n_estimators of the random forest model's hyperparameters. This n_estimators configuration is expected to increase to the optimal values reported in Section 4.1 once full tuning is performed.

5 Discussion

This study investigated the optimization of multi-floor IPS through a systematic evaluation of environmental factors influencing fingerprinting performance when using ML approaches. Our analysis focused on two main factors: the grid size for fingerprint data collection and the impact of low-relevance BSSIDs in the area network. Through extensive experimentation across different grid sizes, several key insights emerged.

First, grid size was found to have a substantial impact on model performance across all evaluation metrics. Contrary to the common assumption that smaller

grid sizes always enhance precision, our findings indicate that a moderate grid size, specifically $7 \times 7\text{m}$, strikes the best balance between spatial resolution and positioning precision. Models trained on this grid size consistently achieved the highest accuracy (approximately 80%), along with the lowest AGT and ADT values. This suggests that while finer grids may provide more data points, they also introduce variability that can overwhelm ML models without delivering proportional gains in localization accuracy.

Second, the analysis of filtering low-relevance BSSIDs showed only marginal improvements. While excluding low-intensity signals (e.g., -100 dBm from distant or low-relevance access points) slightly enhanced accuracy and reduced AGT, the gains were minimal. This indicates that although data cleaning contributes to input quality, the spatial configuration of data—particularly grid structuring—plays a more dominant role in IPS performance for multi-floor environments.

Among the ML models evaluated, RF and XGBoost consistently delivered superior performance across all metrics and grid configurations. These ensemble methods appear well-suited to the fingerprint-matching task, likely due to their ability to capture complex feature interactions and resist over-fitting, particularly in environments with noisy or redundant RSSI signals. Interestingly, more complex deep learning models such as MLP did not outperform these simpler ML techniques, supporting the view that in settings with constrained or moderately sized datasets, traditional models can offer a better trade-off between accuracy and computational efficiency.

It is important to contextualize the generalizability of our findings. The test environment, which is a university campus building with realistic multi-floor usage, ensures ecological validity but certain parameters (e.g., architectural layout, number of access points, and building materials) are inherently site-specific. As such, while the identified optimal grid size and preferred models provide valuable guidance, they may require recalibration for different environments such as warehouses, shopping malls, or hospitals.

Finally, the proposed methodology of using interpolation to synthesize multiple grid sizes from a single fine-grained dataset substantially reduces the labor intensiveness of traditional IPS data collection. This approach not only improves reproducibility and scalability but also establishes a practical framework for future IPS research to systematically explore environmental configurations.

6 Conclusion and Future Directions

This study explored the optimization of multi-floor IPS using grid-based fingerprinting and machine learning. By systematically analyzing the impact of grid size and RSSI values of low-relevance BSSIDs on IPS performance, we identified that a grid size of $7 \times 7\text{m}$ offers the best balance between accuracy and practicality of data collection. Our findings demonstrated that ensemble-based ML models such as RF and XGBoost consistently outperformed other algorithms, achieving high accuracy and low prediction error across multiple evaluation metrics. While

filtering out low-relevance BSSIDs and their RSSI values slightly improved model performance, the overall influence was modest, suggesting that strategic spatial structuring has a more significant effect on accuracy than simple data cleaning.

Looking forward, there are several directions for future research. First, exploring more advanced deep learning models, such as attention-based networks or hybrid GNN-MLP architectures, could uncover further performance gains, especially when trained on larger and more diverse datasets. Additionally, implementing real-time IPS applications with adaptive grid configurations may help tailor accuracy to user needs in dynamic environments. Finally, future work may also examine the use of synthetic data augmentation, sensor fusion, and transfer learning approaches to improve generalization across different buildings or floors.

References

1. Bibbò, L., Carotenuto, R., Della Corte, F.: An overview of indoor localization system for human activity recognition (har) in healthcare. *Sensors* **22**(21) (2022). <https://doi.org/10.3390/s22218119>, <https://www.mdpi.com/1424-8220/22/21/8119>
2. Christodoulou, D.: Developing an Indoor Localization and Wayfinding App for a University Library. Master's thesis, School of Informatics, University of Edinburgh (2022)
3. Csik, D., Odry, A., Sarcevic, P.: Comparison of rssi-based fingerprinting methods for indoor localization. pp. 000273–000278 (09 2022). <https://doi.org/10.1109/SISY56759.2022.10036270>
4. Ezzati Khatab, Z., Moghtadaiee, V., Ghorashi, S.A.: A fingerprint-based technique for indoor localization using fuzzy least squares support vector machine. In: 2017 Iranian Conference on Electrical Engineering (ICEE). pp. 1944–1949 (2017). <https://doi.org/10.1109/IranianCEE.2017.7985373>
5. Freund, Y., Schapire, R.E.: A short introduction to boosting (1999), <https://api.semanticscholar.org/CorpusID:9621074>
6. Gidey, H.T., Guo, X., Zhong, K., Li, L., Zhang, Y.: Ohettla: An online transfer learning method for fingerprint-based indoor positioning. *Sensors* **22**(23) (2022). <https://doi.org/10.3390/s22239044>, <https://www.mdpi.com/1424-8220/22/23/9044>
7. Intachuen, B., Charoenphon, M., Mankhetwit, T.: Classification-based ips (2024), available at: <https://github.com/RinRin-32/Classification-based-IPS>
8. Link, J.A.B., Smith, P., Viol, N., Wehrle, K.: Footpath: Accurate map-based indoor navigation using smartphones. In: 2011 International Conference on Indoor Positioning and Indoor Navigation. pp. 1–8 (2011). <https://doi.org/10.1109/IPIN.2011.6071934>
9. Maung Maung, N.A., Lwi, B.Y., Thida, S.: An enhanced rss fingerprinting-based wireless indoor positioning using random forest classifier. In: 2020 International Conference on Advanced Information Technologies (ICAIT). pp. 59–63 (2020). <https://doi.org/10.1109/ICAIT51105.2020.9261776>
10. Nakpaen, L., Wongsekleo, P., Cherntanomwong, P., Pattiyanon, C.: Building rssi-based indoor positioning fingerprint maps using android-based coordination. In: 2024 19th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP). pp. 1–6 (2024). <https://doi.org/10.1109/iSAI-NLP64410.2024.10799385>

11. Nessa, A., Adhikari, B., Hussain, F., Fernando, X.N.: A survey of machine learning for indoor positioning. *IEEE Access* **8**, 214945–214965 (2020). <https://doi.org/10.1109/ACCESS.2020.3039271>
12. Obeidat, H., Shuaieb, W., Obeidat, O., Abd-Alhameed, R.: A review of indoor localization techniques and wireless technologies. *Wireless Personal Communications* **119** (07 2021). <https://doi.org/10.1007/s11277-021-08209-5>
13. Satan, A.: Bluetooth-based indoor navigation mobile system. In: 2018 19th International Carpathian Control Conference (ICCC). pp. 332–337 (2018). <https://doi.org/10.1109/CarpathianCC.2018.8399651>
14. de Souza Mourão, H.A., de Oliveira, H.A.B.F.: Indoor localization system using fingerprinting and novelty detection for evaluation of confidence. *Future Internet* **14**(2) (2022). <https://doi.org/10.3390/fi14020051>, <https://www.mdpi.com/1999-5903/14/2/51>
15. Vishwakarma, R., Joshi, R.B., Mishra, S.: Indoorgnn: A graph neural network based approach for indoor localization using wifi rssi. In: Goyal, V., Kumar, N., Bhowmick, S.S., Goyal, P., Goyal, N., Kumar, D. (eds.) Big Data and Artificial Intelligence. pp. 150–165. Springer Nature Switzerland, Cham (2023)
16. Wongsekleo, P., Nakpaen, L., Cherntanomwong, P., Pattiyanon, C.: Time reduction for collecting fingerprint data in indoor positioning systems with generated synthetic data by ensemble models and gans. pp. 1–6 (11 2024). <https://doi.org/10.1109/iSAI-NLP64410.2024.10799319>

Appendix

A Implementation Details: Grid Alignment Challenges

During the data collection phase, several practical challenges emerged that required careful consideration.

A.1 Floor Tiling Variations

Different layouts and tiling across areas presented challenges, as certain floors had distinct tiling with varying elevations and angles (as shown in Fig. 11). This was addressed by constructing grids based on previous reference points. Although minor measurement inaccuracies may exist, they are negligible for the purpose of this study.

A.2 Edge Grid Effects

While taping and collecting data points, another factor that could interfere with the IPS collection presented itself. The specific area inside the grid also affected how many RSSI values were present. For instance, collecting data in the top left corner of the grid could produce different RSSI values compared to the center of the grid.

This issue could be addressed by either (1) removing the edge grids and (2) collecting fewer data points in the edge grid. However, both methods result in fewer data points being collected and have minimal impact on IPS performance. This is based on the assumption that, from a user experience perspective, being located at the edge of a grid represents a transitional state between adjacent grids. Meaning that, if misclassification occurs, it typically involves the neighboring grid and the correct grid itself. Therefore, for this study, the edge grid factor is negligible and omitted from further analysis.

B Hardware Implementation Considerations

During the experimentation phase, specific hardware limitations significantly impacted the feasibility of training models with different feature sets. When attempting to train on the full, unfiltered BSSID set (1799 features), computational demands quickly became impractical—the RTX 3080 Ti GPU struggled even with the smallest grid size configurations. Although exact runtimes were not recorded, the difference in resource demands between filtered and unfiltered datasets was substantial.

By applying the simple filtering approach that reduced features from 1799 to 378 BSSIDs, model training became feasible across the entire range of grid sizes (1×1 to 15×15). This hardware constraint, while limiting, provided valuable insights into the practical utility of the full feature set and suggests that systematic feature selection or dimensionality reduction studies could be worthwhile



Fig. 11. An example of taping that was done to create grids for data collection of RSSI values as fingerprints

in future works. The filtering approach not only made training computationally feasible but also demonstrated that selective feature reduction can maintain or even improve model performance while significantly reducing computational requirements.

C Detailed Model Performance Analysis

This section provides comprehensive performance metrics for all machine learning models evaluated in this study across key grid sizes ($1 \times 1\text{m}$, $7 \times 7\text{m}$, and $15 \times 15\text{m}$). The analysis includes accuracy and custom metric (AGT) measurements for filtered BSSID datasets. These detailed results address the comparative performance of kNN, MLP, SVM, XGBoost, and XGBoost Random Forest models.

kNN

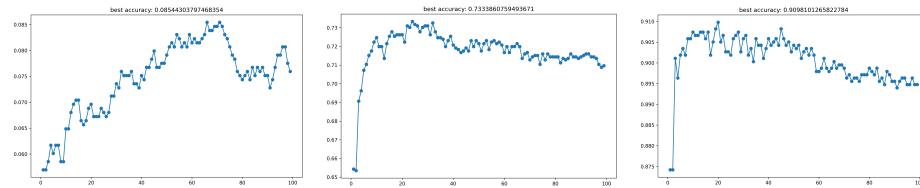


Fig. 12. *
Accuracy: $1 \times 1\text{m}$

Fig. 13. *
Accuracy: $7 \times 7\text{m}$

Fig. 14. *
Accuracy: $15 \times 15\text{m}$

Fig. 15. kNN Accuracy vs Parameter Settings - Grid sizes $1 \times 1\text{m}$, $7 \times 7\text{m}$, and $15 \times 15\text{m}$ (filtered)

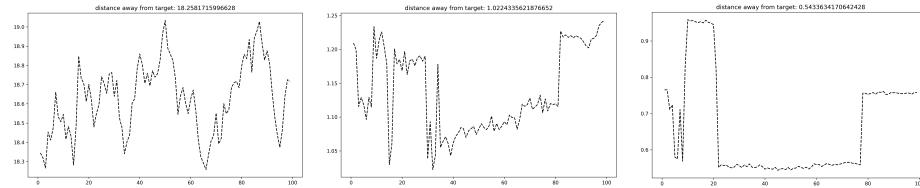


Fig. 16. *
AGT: $1 \times 1\text{m}$

Fig. 17. *
AGT: $7 \times 7\text{m}$

Fig. 18. *
AGT: $15 \times 15\text{m}$

Fig. 19. kNN Custom Metric (AGT) vs Parameter Settings - Grid sizes $1 \times 1\text{m}$, $7 \times 7\text{m}$, and $15 \times 15\text{m}$ (filtered)

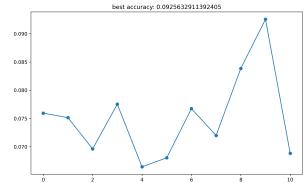
MLP

Fig. 20. *
Accuracy: $1 \times 1\text{m}$

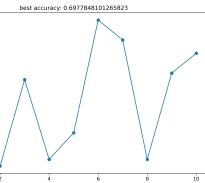


Fig. 21. *
Accuracy: $7 \times 7\text{m}$

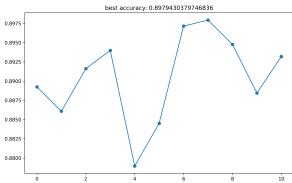


Fig. 22. *
Accuracy: $15 \times 15\text{m}$

Fig. 23. MLP Accuracy vs Parameter Settings - Grid sizes $1 \times 1\text{m}$, $7 \times 7\text{m}$, and $15 \times 15\text{m}$ (filtered)

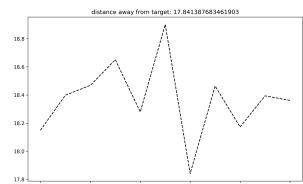


Fig. 24. *
AGT: $1 \times 1\text{m}$

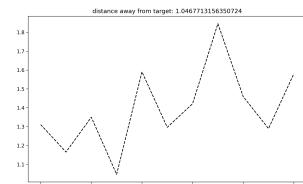


Fig. 25. *
AGT: $7 \times 7\text{m}$

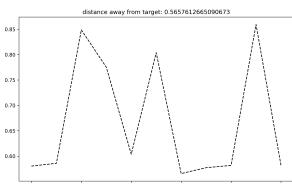


Fig. 26. *
AGT: $15 \times 15\text{m}$

Fig. 27. MLP Custom Metric (AGT) vs Parameter Settings - Grid sizes $1 \times 1\text{m}$, $7 \times 7\text{m}$, and $15 \times 15\text{m}$ (filtered)

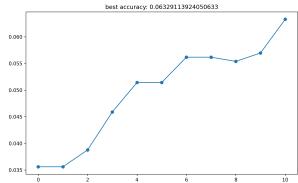
SVM

Fig. 28. *
Accuracy: 1×1m

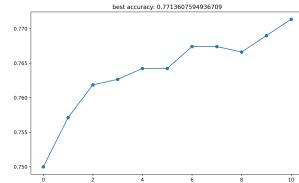


Fig. 29. *
Accuracy: 7×7m

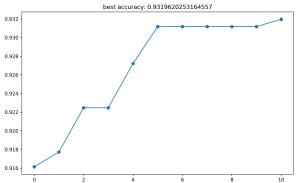


Fig. 30. *
Accuracy: 15×15m

Fig. 31. SVM Accuracy vs Parameter Settings - Grid sizes 1×1m, 7×7m, and 15×15m (filtered)

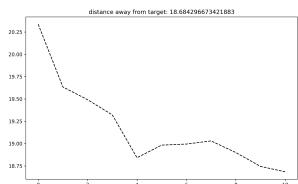


Fig. 32. *
AGT: 1×1m

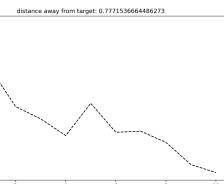


Fig. 33. *
AGT: 7×7m

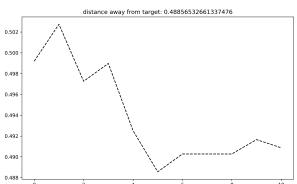


Fig. 34. *
AGT: 15×15m

Fig. 35. SVM Custom Metric (AGT) vs Parameter Settings - Grid sizes 1×1m, 7×7m, and 15×15m (filtered)

XGBoost

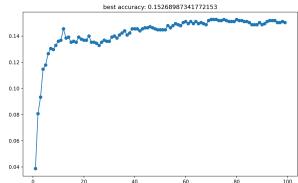


Fig. 36. *
Accuracy: 1×1m

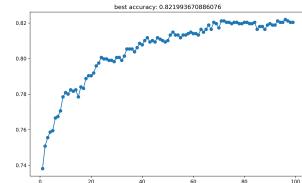


Fig. 37. *
Accuracy: 7×7m

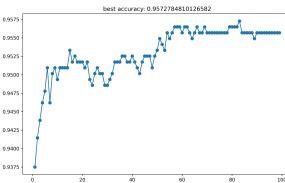


Fig. 38. *
Accuracy: 15×15m

Fig. 39. XGBoost Accuracy vs Parameter Settings - Grid sizes 1×1m, 7×7m, and 15×15m (filtered)

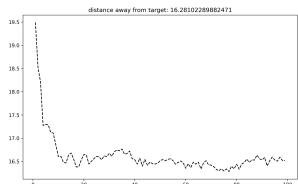


Fig. 40. *
AGT: 1×1m

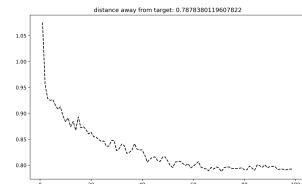


Fig. 41. *
AGT: 7×7m

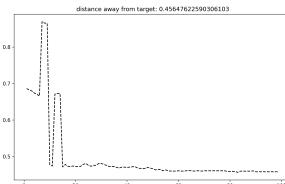


Fig. 42. *
AGT: 15×15m

Fig. 43. XGBoost Custom Metric (AGT) vs Parameter Settings - Grid sizes 1×1m, 7×7m, and 15×15m (filtered)

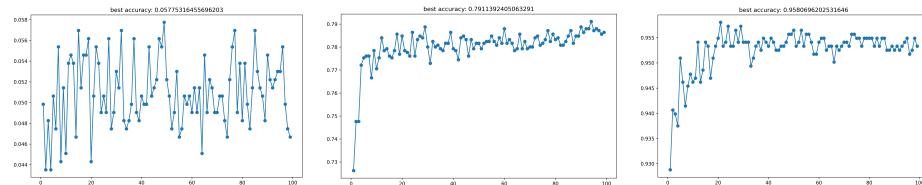
XGBoost Random Forest

Fig. 47. XGBoost Random Forest Accuracy vs Parameter Settings - Grid sizes $1 \times 1\text{m}$, $7 \times 7\text{m}$, and $15 \times 15\text{m}$ (filtered)

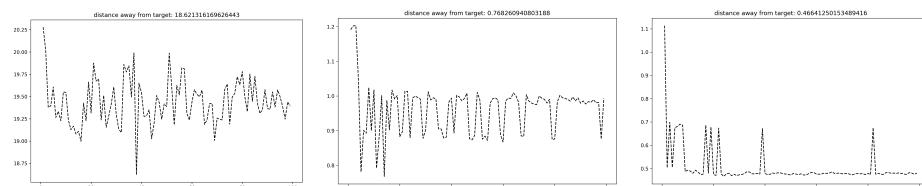


Fig. 51. XGBoost Random Forest Custom Metric (AGT) vs Parameter Settings - Grid sizes $1 \times 1\text{m}$, $7 \times 7\text{m}$, and $15 \times 15\text{m}$ (filtered)