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

Optimizing Grid Size and Feature Selection for ML-Based IPS

Burin Intachuen[1] Mhadhanagul Charoenphon[1] Tanakorn Mankhetwit[1] Charnon Pattiyanon[2]

October 13, 2025

[1]Department of Computer Engineering, Mahidol University (International College), Thailand
[2]Department of AI and Computer Engineering, CMKL University, Thailand

Table of contents

Introduction

Indoor Positioning Systems (IPS)

- GPS and GNSS fail indoors due to signal attenuation
- Building materials (concrete, metal) block satellite signals
- IPS provides alternative solution for indoor navigation
- Applications: shopping malls, hospitals, university campuses

Challenge
How to optimize IPS performance in multi-floor environments?

Current Approaches

Traditional Methods:

- Trilateration using BLE or Wi-Fi RSSI
- Fingerprinting-based positioning
- Limited by environmental noise

Machine Learning Approaches:

- kNN, Random Forest, SVM, Neural Networks
- Treat positioning as classification problem
- Better handling of noisy RSSI data

Research Motivation

Previous work limitations:

- Large grid sizes (16.75×15 m) used for simplicity
- Limited understanding of grid size impact
- No systematic study of feature filtering effects

Our contributions:

1. New evaluation metrics: AGT and ADT
2. Grid size vs. precision trade-off analysis
3. Feature filtering impact on model performance
4. Practical insights for IPS implementation

Research Methodology

Research Overview

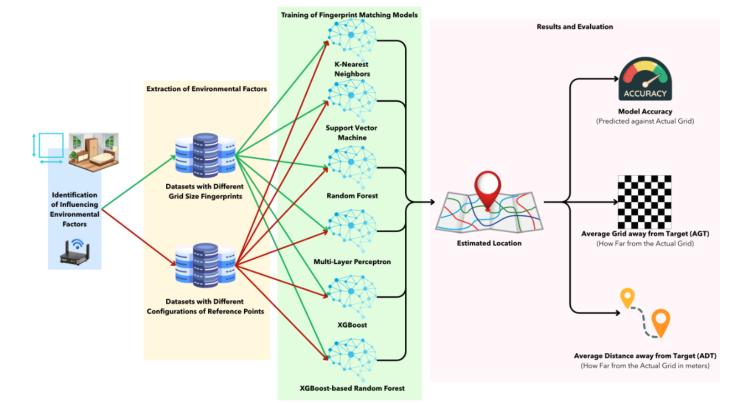

Figure 1: Four-part research methodology

Environmental Factors

Key factors identified:

1. **Grid Size**
    - Larger grids $\rightarrow$ easier classification, lower precision
    - Smaller grids $\rightarrow$ higher precision, more data collection effort
2. **Low-Relevance BSSIDs**
    - Signals from distant access points
    - Introduce noise and increase computational cost
    - Need systematic filtering approach

RSSI Heatmap Analysis

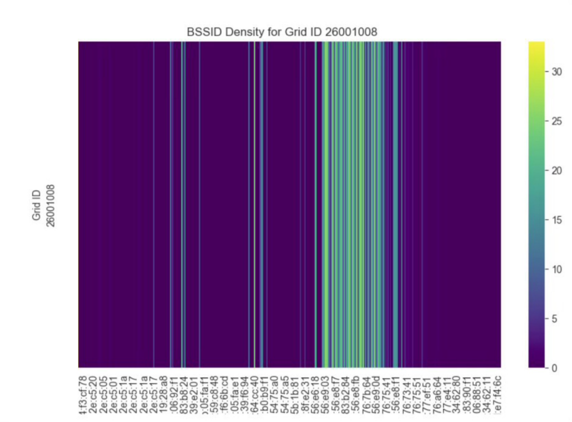

Figure 2: Heatmap showing RSSI values from different access points

- Darker colors = lower signal strength
- Many low-relevance BSSIDs detected

New Evaluation Metrics

Average Grid from Target (AGT):

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

Average Distance from Target (ADT):

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

- AGT: Grid-based distance (adaptable to any grid size)
- ADT: Physical distance in meters

Grid Interpolation Approach

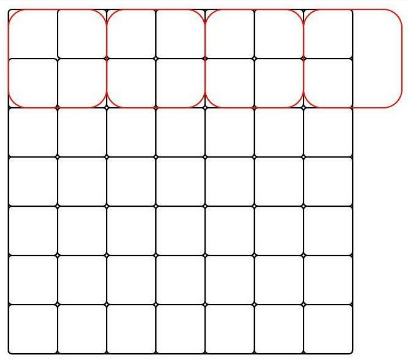

Figure 3: Grid aggregation visualization

Benefits:

- Avoid re-taping entire area
- Generate multiple grid sizes from 1×1m base
- Reduce human measurement errors

Method:

- Average RSSI from 5 sampling points
- Test 8 grid sizes: 1×1m to 15×15m

Dataset Overview

Table 1: Dataset metadata by grid size

Floor	Grid Size	# Grids	# Data Points
1st	1×1m	309	8,086
1st	7×7m	17	445
1st	15×15m	5	130
2nd	1×1m	174	4,554
2nd	7×7m	12	314
2nd	15×15m	4	105

- Total: 12,640 RSSI samples
- Filtered: 378 BSSIDs (from 1,799)

Machine Learning Models

Models evaluated:

- k-Nearest Neighbor (kNN)
- Random Forest (RF)
- Support Vector Machine (SVM)
- Multi-Layer Perceptron (MLP)
- XGBoost
- XGBoost + Random Forest hybrid

Training approach:

- Standard process with hyperparameter tuning
- Consistent dataset across all models
- Comparison of filtered vs. unfiltered BSSIDs

Experiments and Results

Accuracy by Grid Size

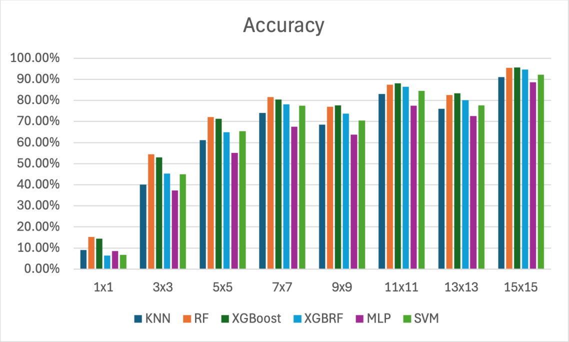

Figure 4: Model accuracy across different grid sizes (Filtered)

- Larger grids $\rightarrow$ higher accuracy
- 7×7m grid achieves 80% accuracy
- RF and XGBoost perform best

AGT Results

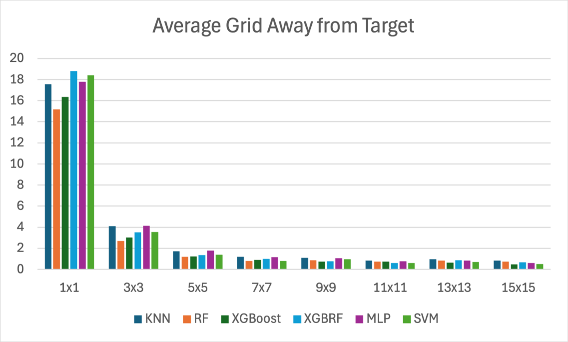

Figure 5: Average Grid from Target (Filtered)

- AGT stabilizes at 7×7m grid size
- Smaller grids show higher positioning errors

ADT Results

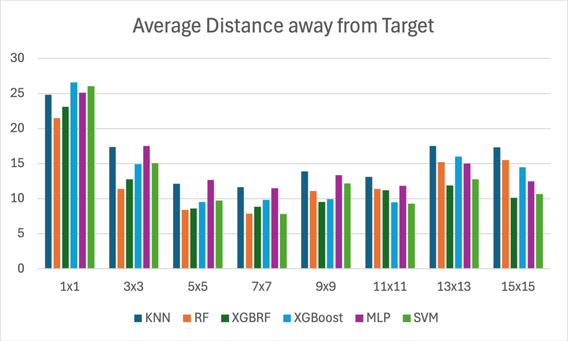

Figure 6: Average Distance from Target in meters (Filtered)

- 7×7m grid minimizes physical distance error
- Optimal balance for precision-dependent applications

Feature Filtering Impact

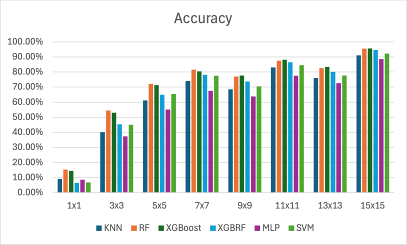

Figure 7: Filtered (378 BSSIDs)

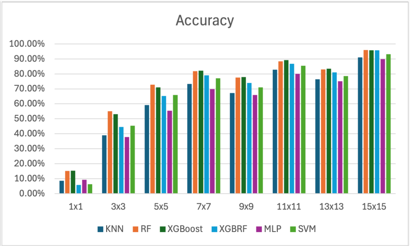

Figure 8: Unfiltered (1,799 BSSIDs)

Filtering benefits:

- Comparable accuracy with 79% fewer features
- More stable performance
- Reduced computational requirements

Discussion

Key Findings

1. **Optimal Grid Size: 7×7m**
   - Best balance between accuracy and spatial resolution
   - 80% accuracy with lowest AGT/ADT values
2. **Feature Filtering Works**
   - 79% reduction in features (1,799 → 378)
   - Maintained performance, improved stability
   - Significantly reduced training time
3. **Best Models: RF and XGBoost**
   - Consistently outperformed other models
   - Well-suited for fingerprint-matching tasks
   - Better than complex deep learning in this setting

Practical Implications

For IPS Implementation:

- Don't assume smaller grids are always better
- Moderate grid sizes can be optimal
- Simple feature filtering is effective
- Traditional ML can outperform deep learning

Limitations:

- Site-specific results (university campus)
- May need recalibration for different environments
- Architecture and materials affect performance

Conclusion

Summary

What we achieved:

- Identified optimal 7×7m grid size for multi-floor IPS
- Introduced AGT and ADT metrics for standardized evaluation
- Demonstrated effective BSSID filtering approach
- Showed RF and XGBoost superiority

Future directions:

- Advanced deep learning models (attention-based, GNN-MLP)
- Real-time adaptive grid configurations
- Synthetic data augmentation and sensor fusion
- Transfer learning across buildings

Questions?

Implementation Challenges

Floor tiling variations:

- Different elevations and angles
- Addressed by reference-point-based grid construction

Edge grid effects:

- Corner vs. center RSSI variations
- Minimal impact on overall performance
- Misclassifications typically involve adjacent grids

Hardware constraints:

- RTX 3080 Ti struggled with unfiltered dataset
- Filtering made training feasible across all grid sizes