RSSI-Based Passenger Movement Classification for Non-Intrusive Public Transport Monitoring

Author One*, Author Two†, Author Three‡ and Author Four§
Department of Whatever, Whichever University
Wherever
Email: *author.one@add.on.net, †author.two@add.on.net, ‡author.three@add.on.net, §author.four@add.on.net

***Abstract*—Accurate monitoring of passenger flows in public transport is essential for service optimization and network planning, yet traditional counting methods face limitations in coverage, cost, and privacy preservation. This paper presents a novel approach for classifying passenger movements using temporal sequences of WiFi Received Signal Strength Indicator (RSSI) measurements. We introduce a dataset collected in a controlled experimental environment that simulates public transport scenarios, capturing the distinctive signal patterns associated with four fundamental movement classes: boarding the vehicle, alighting from the vehicle, remaining inside, and remaining at the bus stop. By analyzing the temporal evolution of RSSI values over ten-second observation windows, our approach enables non-intrusive distinction between static and transitional states without requiring specialized hardware or compromising passenger anonymity. Experimental evaluation using multiple machine learning classifiers demonstrates the feasibility of RSSI-based movement classification, providing a cost-effective complement to existing automatic passenger counting systems for intelligent transportation applications.**

***Index Terms*—Passenger counting, RSSI fingerprinting, WiFi sensing, public transport, machine learning, intelligent transportation systems, urban mobility**

I. Introduction

Urban mobility remains a pressing concern, as over 75% of EU citizens reside in cities and transportation accounts for roughly 24% of greenhouse gas emissions [1], [2]. Managing public transport effectively hinges on accurate passenger flow data, yet traditional counting approaches—manual surveys, ticketing systems, infrared sensors, and automated passenger counting (APC) devices—come with drawbacks such as high costs, incomplete coverage, and privacy concerns [3], [4]. Moreover, conventional ticketing often fails to capture complete passenger journeys when exit validation is absent [5].

The widespread adoption of WiFi-enabled devices opens up new avenues for non-intrusive passenger sensing. Such devices generate Received Signal Strength Indicator (RSSI) signatures that allow movement pattern recognition while preserving anonymity [6], [7]. Although Channel State Information (CSI) approaches can achieve high accuracy, they demand specialized hardware and intensive computation [8]. RSSI-based methods, by contrast, offer a practical alternative using standard networking equipment [9].

This paper puts forward a new approach to passenger boarding and alighting classification using temporal RSSI sequences. By tracking signal evolution over ten-second observation windows, our methodology distinguishes four movement patterns: remaining inside the vehicle, remaining at the stop, boarding, and alighting—all without requiring precise localization.

The principal contributions of this work are: (1) an RSSI-based movement classification framework that exploits temporal signal evolution; (2) an experimental dataset comprising approximately 1,360 labelled samples across four classes; (3) a thorough evaluation of 38 machine learning classifiers; (4) feature importance analysis; and (5) a privacy-preserving sensing approach that operates without device identification.

The remainder of this paper reviews related work (section II), describes the experimental setup (section III), presents exploratory data analysis (section IV), presents results (section V) and discussion (section VI), and concludes with future directions (section VIII).

II. Related Work

A. Automatic Passenger Counting Systems

Conventional APC systems rely on infrared sensors, pressure mats, or video-based detection [3]. Pronello and Garzón Ruiz [4] observed that claimed 98% accuracy frequently drops to 53–74% under real-world conditions. Deep learning approaches can reach up to 94% accuracy [10], though vision-based systems remain hampered by occlusion, lighting variability, and privacy concerns.

B. WiFi-Based Passenger Sensing

The ubiquity of WiFi-enabled devices has spurred interest in wireless signal analysis for mobility monitoring. Myrvoll et al. [11] pioneered probe request analysis for passenger counting, while Nitti et al. [6] reported 100% detection in static settings and 94% in dynamic ones. CSI-based systems yield richer information, with Guo et al. [8] reaching over 94% accuracy, though they require specialized hardware. RSSI remains practical with off-the-shelf equipment. Fabre et al. [12] found Light Gradient Boosting effective for WiFi-based ridership estimation, and Simončič et al. [7] attained over 96% accuracy despite MAC randomization.

C. RSSI Fingerprinting and Movement Classification

RSSI fingerprinting is well-established for indoor localization [9]. Wang et al. [13] proposed treating RSSI as temporal sequences, aligning with our methodology. Servizi et al. [14]

addressed boarding detection using Bluetooth sensing, while Cerqueira et al. [5] demonstrated the importance of complete journey patterns for OD matrix inference.

D. Research Gap

Despite progress in passenger counting and wireless sensing, machine learning classification of movement patterns from RSSI time series has received comparatively little attention. Existing approaches tend to focus on aggregate counting rather than fine-grained movement classification. Our work addresses this gap by casting passenger movement detection as a supervised classification problem using temporal RSSI evolution, thereby enabling real-time trajectory classification with a single access point at the vehicle door and keeping infrastructure requirements to a minimum.

III. Experimental Setup

This section outlines the physical data collection environment and the machine learning experimental framework devised for passenger movement classification.

A. Physical Data Collection Setup

Controlled experiments were carried out in an indoor environment that emulates public transport interactions, allowing reproducible data collection under both isolated and noisy conditions.

1) Environmental Configuration: The environment was divided into two zones (Figure 1). Zone A (Vehicle Interior) consisted of a closed room simulating a bus interior, with a WiFi access point positioned adjacent to the doorway. Zone B (Bus Stop) was the corridor outside, representing the boarding area. The wall and door between zones introduce signal attenuation, generating distinctive RSSI patterns during transitions.

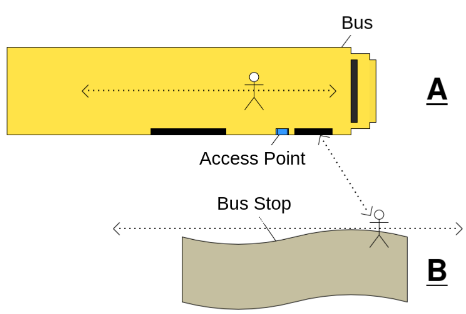

Figure 1. Experimental environment simulating a public transport scenario.

2) Data Acquisition: Data was collected using a Python script interfacing with the AP via Ethernet. Each trial comprised a **10-second window** with **10 RSSI samples at 1 Hz**. Devices maintained periodic low-overhead traffic (ICMP) to ensure consistent RSSI reporting. Four mobile devices from three manufacturers were used to introduce hardware variability.

Four movement classes were defined: **A→A** (remaining inside), **B→B** (remaining at stop), **A→B** (alighting), and **B→A** (boarding). Collection occurred under two scenarios: *Isolated* (single device, 20 repetitions per class per device) and *Noisy* (four devices simultaneously performing paired movements).

3) Preprocessing and Dataset Structure: Raw data was transformed through: (1) *Temporal Aggregation*—10 RSSI measurements per trial aggregated into feature vector $\mathbf{R} = [r_1, \ldots, r_{10}]$; (2) *Device Isolation* using MAC addresses; (3) *Labeling* with movement class and noise indicator; (4) *Feature Filtering* to retain only RSSI values. The resulting CSV dataset contains approximately 1,356 samples, each representing a 10-second trajectory with movement class label, enabling analysis of both absolute signal strength and temporal evolution.

4) Experimental Scenarios: To comprehensively evaluate the impact of data collection conditions on classification performance, three distinct experimental scenarios were defined:

Combined Dataset: The complete dataset containing all 1,356 samples from both isolated and noisy collection conditions. This scenario represents the most realistic deployment setting, where the classifier must generalize across varying environmental conditions.

Isolated-Only Dataset: A subset containing exclusively samples collected under isolated conditions (single device, $n = 160$). This scenario provides an upper-bound estimate of classification performance under ideal conditions with minimal signal interference.

Noisy-Only Dataset: A subset containing exclusively samples collected with simultaneous multi-device activity ($n = 1,196$). This scenario evaluates classifier robustness under challenging conditions that closely approximate real-world public transport environments.

Comparing across these scenarios allows for a quantitative assessment of how environmental noise affects classification accuracy and sheds light on the operational boundaries of RSSI-based movement detection.

B. Machine Learning Experimental Framework

The following subsection details the classifier selection rationale, evaluation methodology, and performance metrics adopted in our experimental protocol.

1) Classifier Selection and Justification: A total of 38 classification algorithms spanning multiple paradigms were evaluated to ensure thorough benchmarking. The classifier families were chosen based on their established effectiveness in RSSI-based classification tasks [15], [16]:

Support Vector Machines (SVM): SVMs with RBF and linear kernels were included due to their demonstrated superiority in WiFi fingerprinting tasks. Prior studies on indoor localization using RSSI have shown SVMs achieving accuracies exceeding 90% for location classification [9], [17]. The RBF kernel effectively captures non-linear relationships in signal strength patterns.

Ensemble Methods: Random Forest and Extra Trees were selected for their robustness to noise and ability to model complex decision boundaries without extensive hyperparameter tuning [15]. Gradient boosting variants (XGBoost, LightGBM, CatBoost) were included based on their state-of-the-art performance in tabular classification tasks, with CatBoost demonstrating particular effectiveness for categorical features [12].

Gaussian Process Classifier: GPs provide probabilistic predictions with uncertainty quantification, particularly valuable for RSSI data where signal variability is inherent. The RBF kernel enables automatic adaptation to the intrinsic dimensionality of temporal RSSI sequences.

Neural Networks: Multi-layer perceptrons (MLPs) with varying architectures (small, medium, large) were evaluated to assess whether deeper representations improve classification over traditional methods for this feature space dimensionality.

Stacking and Voting Ensembles: Meta-learning approaches combining heterogeneous base learners were included to leverage complementary classifier strengths, a strategy shown to improve robustness in transportation sensing applications [10].

2) Data Partitioning Strategy: The dataset was partitioned using **stratified sampling** with an 80%/20% train-test split. Stratified sampling ensures that class distributions are preserved in both partitions, which is critical for multi-class classification problems where class imbalance could otherwise bias model evaluation. This approach maintains the original proportion of each movement class (AA, BB, AB, BA) in both training and testing sets.

3) Cross-Validation Protocol: Model training employed **5-fold stratified cross-validation**, a methodology widely recommended for robust classifier evaluation. Stratified K-fold cross-validation maintains class ratios across all folds, ensuring that minority classes receive adequate representation during training and validation. This technique reduces variance in performance estimates compared to simple hold-out validation.

To assess result stability, experiments were repeated with three random seeds (3, 5, and 42), and metrics were aggregated across runs. This multi-seed evaluation quantifies classifier sensitivity to random initialization and data shuffling.

4) Performance Metrics: Four complementary metrics were employed to provide comprehensive performance characterization:

Accuracy: The proportion of correctly classified samples. While intuitive, accuracy can be misleading for imbalanced datasets.

Weighted F1-Score: The harmonic mean of precision and recall, weighted by class support. This metric balances false positives and false negatives while accounting for class distribution.

Balanced Accuracy: The arithmetic mean of per-class recall values, ensuring equal contribution from each class regardless of prevalence.

Matthews Correlation Coefficient (MCC): Selected as the primary evaluation criterion, as this metric offers a reliable and balanced evaluation of classification models, particularly in scenarios involving imbalanced datasets or when assessing performance across multiple classes [18], [19]. MCC ranges from -1 to $+1$, producing high scores only when all confusion matrix categories achieve strong results.

IV. Exploratory Data Analysis

Ahead of classifier training, an exploratory data analysis was carried out to gauge the discriminative potential of temporal RSSI signatures and to characterize signal behaviour across different passenger movement classes.

A. Dataset Composition

The final dataset is roughly balanced, comprising around 340 samples per movement class. For each class, 40 samples were gathered under isolated conditions, while the remaining samples were obtained during simultaneous device activity, thereby introducing controlled signal interference. This balance ensures that the exploratory analysis and subsequent results reflect both ideal and realistic operating conditions.

B. Temporal RSSI Characteristics

The temporal evolution of RSSI values serves as the primary discriminative feature between movement classes. Figure 2 depicts the mean RSSI trajectory over the 10-second observation window for each class.

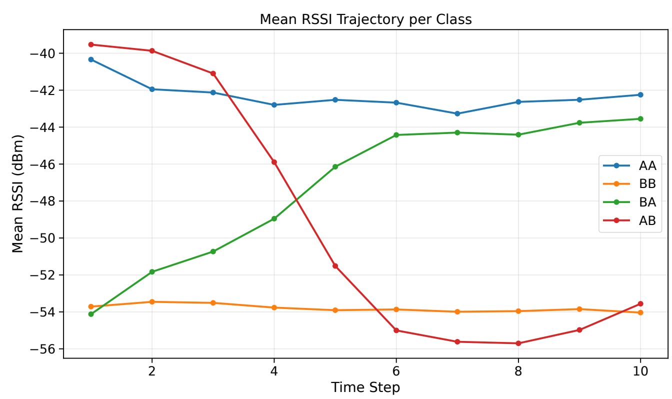

Figure 2. Temporal evolution of RSSI values over the 10-second observation window for each movement class.

Static states (**AA and BB**) exhibit relatively stable signal levels over time, albeit with distinct average magnitudes due to their spatial separation from the access point. In contrast, transitional movements display clear monotonic trends. The boarding class (**B** $\rightarrow$ **A**) shows a consistent increase in RSSI as the devices move toward the access point, while the alighting class (**A** $\rightarrow$ **B**) presents a pronounced decrease as physical obstructions attenuate the signal.

These contrasting temporal patterns offer a compelling rationale for leveraging RSSI sequences in movement classification.

C. Feature Space Separability

To further inspect the structure of the 10-dimensional RSSI feature vectors, a t-Distributed Stochastic Neighbor Embedding (t-SNE) projection was applied for visualization purposes. The resulting two-dimensional embedding, shown in Figure 3, reveals the formation of four predominantly distinct clusters corresponding to the defined movement classes.

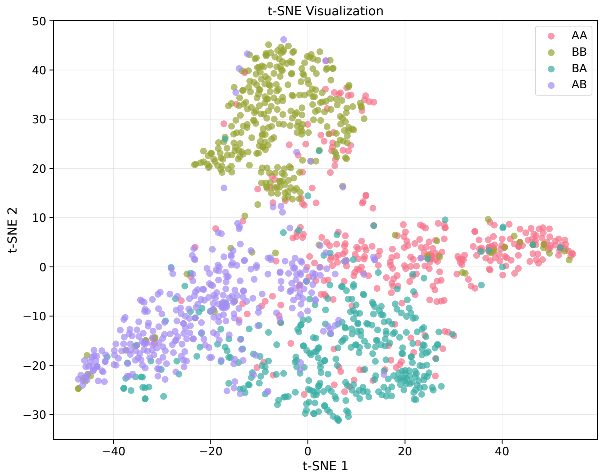

Figure 3. t-SNE projection for visualization of each class structure

Although partial overlap is evident, mainly among samples collected under noisy conditions, the overall clustering suggests that temporal RSSI patterns retain sufficient class-dependent structure to support supervised learning approaches. This analysis serves as a qualitative inspection of feature separability rather than a quantitative performance evaluation.

V. Results

This section presents the experimental results from training and evaluating an extensive set of machine learning classifiers on the RSSI-based passenger movement dataset across three experimental scenarios: combined, isolated-only, and noisy-only conditions. The evaluation covers 38 classification algorithms assessed across multiple random seeds to ensure statistical robustness.

A. Classification Performance: Combined Dataset

Table I summarizes the performance of the top-10 classifiers on the combined dataset, ranked by mean MCC across the experimental seeds.

Table I
Top-10 Classifiers on Combined Dataset (Ranked by MCC)

Classifier	Accuracy	Recall	F1-Score	MCC
GaussianProcess	0.816 ± 0.024	0.816 ± 0.024	0.815 ± 0.024	0.756 ± 0.033
SVC (RBF)	0.815 ± 0.015	0.815 ± 0.015	0.813 ± 0.014	0.755 ± 0.021
SVC (Linear)	0.813 ± 0.017	0.813 ± 0.017	0.812 ± 0.016	0.750 ± 0.023
StackingEnsemble	0.811 ± 0.015	0.811 ± 0.015	0.810 ± 0.014	0.749 ± 0.020
CatBoost	0.809 ± 0.013	0.809 ± 0.013	0.809 ± 0.013	0.746 ± 0.017
RandomForest	0.808 ± 0.017	0.808 ± 0.017	0.807 ± 0.016	0.744 ± 0.022
LogisticRegression (L1)	0.808 ± 0.031	0.808 ± 0.031	0.806 ± 0.031	0.744 ± 0.042
LogisticRegression (ElasticNet)	0.808 ± 0.031	0.808 ± 0.031	0.806 ± 0.031	0.744 ± 0.042
MLP (Large)	0.806 ± 0.015	0.806 ± 0.015	0.805 ± 0.014	0.743 ± 0.021
LogisticRegression (L2)	0.806 ± 0.033	0.806 ± 0.033	0.805 ± 0.033	0.743 ± 0.044

The Gaussian Process classifier attained the highest mean MCC (0.756) with an accuracy of 81.6%. Support Vector Machines with RBF and linear kernels followed closely, recording MCC values of 0.755 and 0.750, respectively.

B. Comparative Analysis Across Experimental Scenarios

Table II presents the performance comparison of top classifiers across the three experimental scenarios, revealing the substantial impact of data collection conditions on classification performance.

Table II
Performance Comparison Across Experimental Scenarios (MCC)

Classifier	Combined	Isolated-Only	Noisy-Only
KNN (k=5)	0.692 ± 0.029	**0.907 ± 0.061**	0.704 ± 0.025
KNN (k=3)	0.690 ± 0.046	0.882 ± 0.068	0.702 ± 0.043
LinearSVC	0.726 ± 0.039	0.867 ± 0.060	0.716 ± 0.031
LogisticRegression (L2)	0.743 ± 0.044	0.866 ± 0.028	0.731 ± 0.009
SVC (Linear)	0.750 ± 0.023	0.850 ± 0.088	0.731 ± 0.037
StackingEnsemble	0.749 ± 0.020	0.851 ± 0.122	0.768 ± 0.023
ExtraTrees	0.737 ± 0.013	0.836 ± 0.041	0.755 ± 0.021
GaussianProcess	**0.756 ± 0.033**	0.414 ± 0.052	0.755 ± 0.028
SVC (RBF)	0.755 ± 0.021	0.825 ± 0.101	0.754 ± 0.016
CatBoost	0.746 ± 0.017	0.782 ± 0.063	**0.770 ± 0.013**

The isolated-only scenario yielded the highest classification performance, with KNN (k=5) reaching an MCC of 0.907—a 20% improvement over the combined dataset. Simpler classifiers such as KNN showed the most pronounced performance gains in the isolated scenario.

The noisy-only scenario exhibited performance levels comparable to the combined dataset, with CatBoost attaining the highest MCC of 0.770. The Gaussian Process classifier maintained robust performance (MCC: 0.755) across both combined and noisy-only conditions, yet showed notable degradation in the isolated scenario (MCC: 0.414).

C. Per-Class Analysis

Table III presents the per-class accuracy, recall, F1-score, and MCC for the best classifier (Gaussian Process).

Table III
Per-Class Performance Metrics (Gaussian Process)

Class	Accuracy	Recall	F1-Score	MCC
AA (Inside)	0.838	0.750	0.791	0.785
BB (Stop)	0.838	0.897	0.878	0.785
BA (Boarding)	0.838	0.824	0.855	0.785
AB (Alighting)	0.838	0.882	0.828	0.785
Weighted Avg	**0.838**	**0.838**	**0.838**	**0.785**

The static state at the bus stop (BB) exhibited the highest recall (89.7%) and F1-score (0.878). The boarding movement (BA) achieved recall of 82.4%, F1-score of 0.855, and MCC of 0.785. The alighting class (AB) demonstrated high recall (88.2%) but lower F1-score (0.828). The static state inside the vehicle (AA) presented the lowest recall (75.0%) and F1-score (0.791). The overall accuracy reached 83.8% with an MCC of 0.785.

D. Hyperparameter Configuration

The Gaussian Process classifier employed a Radial Basis Function (RBF) kernel, which is well-suited for capturing non-linear relationships in the RSSI feature space. The model configuration used is presented in Table IV.

Table IV
GAUSSIAN PROCESS HYPERPARAMETERS

Parameter	**Value**
Kernel	$1.0 \times \mathrm{RBF}(1.0)$
Kernel Length Scale	Optimized during fitting
Optimizer	L-BFGS-B
Max Iterations	100
Multi-class Strategy	One-vs-Rest

The RBF kernel automatically learns the optimal length scale parameter during training, adapting to the intrinsic dimensionality of the RSSI temporal sequences.

E. Confusion Matrix Analysis

Figure 4 presents the normalized confusion matrix for the Gaussian Process classifier, which achieved the best average performance across experimental seeds.

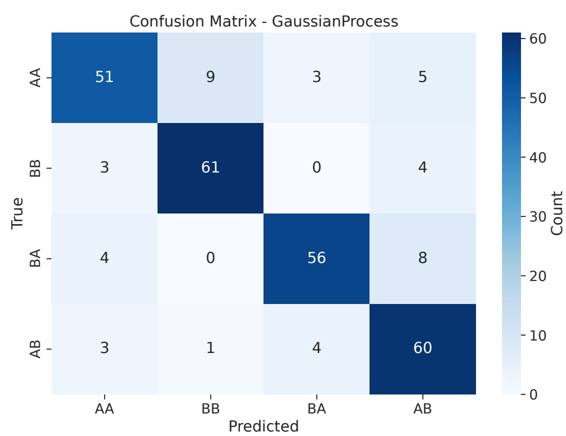

Figure 4. Confusion matrix for the Gaussian Process classifier, demonstrating strong diagonal dominance with minimal inter-class confusion.

The confusion matrix reveals that the primary source of classification errors occurs between spatially adjacent classes. The AA class (remaining inside) is occasionally misclassified as BA (boarding), and minor confusion exists between AB (alighting) and BB (remaining at stop).

F. Model Stability Analysis

To evaluate the robustness of classifier rankings across different experimental conditions, Figure 5 illustrates the accuracy variability for the top classifiers across the three random seeds.

The analysis confirms that top-performing classifiers maintain consistent relative rankings across seeds, with kernel-based methods (Gaussian Process, SVC) and ensemble approaches (Stacking, CatBoost) displaying the lowest variability.

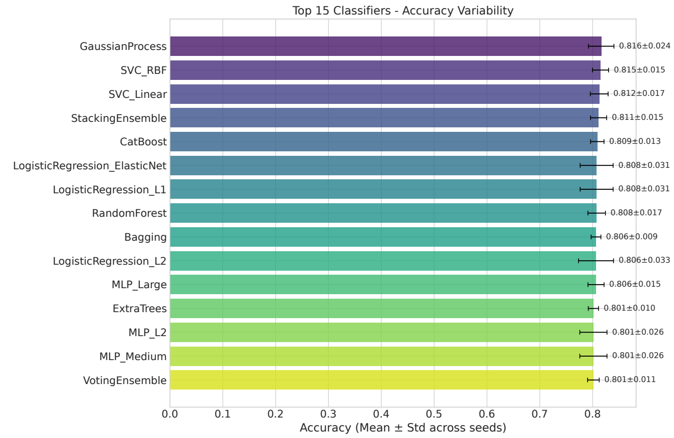

Figure 5. Accuracy variability across experimental seeds for top classifiers, demonstrating consistent ranking stability.

G. Feature Importance

Analysis of feature importance across interpretable classifiers revealed that the initial RSSI measurements (features 1–3) contribute most significantly to classification decisions, as illustrated in Figure 6.

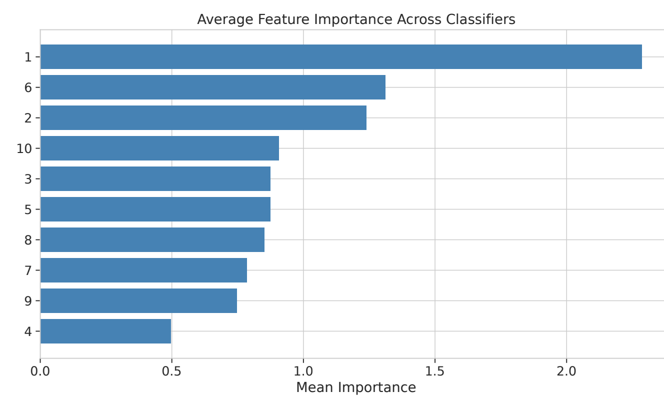

Figure 6. Mean feature importance across classifiers, highlighting the discriminative value of initial RSSI measurements.

The first RSSI sample exhibits the highest importance (normalized score: 2.29), followed by samples at positions 6 and 2.

VI. DISCUSSION

The aforementioned experimental results confirm the feasibility of using temporal RSSI sequences for non-intrusive passenger movement classification in public transport scenarios. This section analyses the key findings across the three experimental scenarios, examines how data collection conditions influence performance, and discusses the practical implications for real-world deployment.

A. Classifier Performance Analysis

The comparative evaluation across experimental scenarios paints a nuanced picture of classifier suitability for RSSI-based movement classification. On the combined dataset, the

Gaussian Process classifier attained the highest MCC (0.756), exhibiting strong discriminative capability for the temporal RSSI patterns. This performance stems from its probabilistic framework and the flexibility of the RBF kernel in modelling non-linear decision boundaries. Unlike parametric models that assume specific functional forms, Gaussian Processes adapt their complexity to the underlying data distribution, an advantage for RSSI patterns that exhibit complex spatial dependencies.

The flexibility of the RBF kernel enables the Gaussian Process to model complex decision boundaries without requiring extensive manual hyperparameter tuning, making it particularly suitable for RSSI-based classification where signal patterns exhibit non-linear spatial dependencies. Support Vector Machines with RBF and linear kernels achieved comparable performance on the combined dataset, confirming that kernel-based methods are well-suited for this classification task. The margin-maximization principle of SVMs provides robust generalization, particularly relevant given the overlap between movement classes observed in the t-SNE visualization.

B. Impact of Data Collection Conditions

Perhaps the most striking finding emerges from the comparison between experimental scenarios. The isolated-only dataset yielded markedly superior classification performance, with KNN (k=5) reaching an MCC of 0.907—a 31% relative improvement over its performance on the combined dataset (MCC: 0.692). This pronounced increase can be traced to the absence of inter-device signal interference during data collection, resulting in cleaner RSSI patterns with more distinct class separations. Two primary factors account for this enhancement:

Signal Clarity: In isolated conditions, RSSI measurements are unaffected by inter-device interference, co-channel contention, or access point load variations. The resulting signal patterns exhibit clearer temporal trajectories with reduced variance within each movement class.

Class Separability: The absence of noise enables more distinct decision boundaries between movement classes. Simpler classifiers such as KNN, which rely on local neighbourhood structure, benefit disproportionately from this increased separability, suggesting that the underlying class boundaries are well-defined when noise is absent. This explains the pronounced performance gains observed for KNN in the isolated scenario.

Conversely, complex models such as the Gaussian Process suffered marked performance degradation in the isolated scenario (MCC: 0.414), despite attaining the best results on the combined dataset. This apparent paradox is explained by the limited sample size ($n = 159$) of the isolated dataset, which proves insufficient for the Gaussian Process to reliably estimate its kernel hyperparameters without overfitting.

The noisy-only scenario, representing the most realistic operational conditions, showed performance levels on par with the combined dataset. CatBoost attained the highest MCC (0.770) in this scenario, suggesting that gradient boosting methods are particularly robust to signal interference. The consistency between noisy-only and combined results indicates that the combined dataset's performance is largely driven by the noisy samples, which make up 88% of the total data.

C. Model Stability Considerations

The observed stability of classifier rankings across experimental seeds has notable implications for deployment. Kernel-based methods (Gaussian Process, SVC) and ensemble approaches (Stacking, CatBoost) showed the lowest variability, a desirable property for deployment scenarios where model retraining may occur with different data partitions. This consistency lends confidence that selected classifiers will maintain their relative performance across varying operational conditions.

D. Classifier Selection Guidelines

The experimental results inform practical recommendations for classifier selection based on deployment context:

High-interference environments: For scenarios with multiple simultaneous devices, gradient boosting methods (CatBoost, XGBoost) and ensemble approaches (StackingEnsemble) offer the best balance of accuracy and robustness.

Controlled environments: In settings with minimal device density, simpler classifiers such as KNN or logistic regression can achieve superior performance while offering reduced computational overhead and improved interpretability.

General deployment: For systems that must operate across varying conditions, SVC with RBF kernel provides consistent performance with acceptable variance, making it a reliable default choice.

E. Per-Class Error Analysis

The confusion patterns reveal insights into the physical characteristics of each movement class. The static state inside the vehicle (AA) exhibited the lowest recall (75.0%), primarily due to misclassification as boarding (BA). This confusion is attributable to the spatial proximity of both classes to the access point, resulting in similar high-RSSI signatures that reflect the similarity in RSSI magnitude when devices are positioned near the access point. Although the temporal dynamics differ (AA maintains relatively stable values while BA shows an increasing trend), this distinction may be subtle in short observation windows.

Conversely, the bus stop class (BB) achieved the highest recall (89.7%), attributable to the consistent low RSSI values observed when devices remain outside the vehicle. The physical barrier separating Zone B from the access point provides natural signal attenuation that enables clear class discrimination. The transitional classes (AB and BA) benefited from their characteristic monotonic RSSI trends, with alighting (AB) achieving 88.2% recall. However, the lower F1-score (0.828) for the alighting class indicates some false positives from the AA class. The boarding movement (BA) achieved strong results (F1-score: 0.855), benefiting from the distinctive increasing RSSI pattern as devices approach the access point.

The overall accuracy of 83.8% and MCC of 0.785 confirm robust multi-class discrimination across all movement categories, demonstrating that RSSI-based temporal patterns provide sufficient information for reliable passenger movement classification.

F. Confusion Matrix Interpretation

The observed confusion between spatially adjacent classes provides important insights for system design. The AA–BA confusion reflects the similarity in RSSI magnitude when devices are positioned near the access point, suggesting that additional features or longer observation windows might improve discrimination between static and transitional states. Similarly, the AB–BB confusion arises because both classes share lower RSSI values characteristic of the exterior zone, though the transitional nature of AB provides some discriminative information through temporal patterns.

G. Feature Importance Insights

The feature importance analysis revealed that initial RSSI measurements (samples 1–3) contribute most significantly to classification decisions, capturing the starting position and providing immediate context for distinguishing static states from transitional movements. This finding aligns with the temporal dynamics of passenger movements, where early signal readings capture the initial position before any state transition occurs.

The elevated importance of sample 6 (mid-trajectory) indicates that classifiers also rely on signal evolution to confirm movement direction, validating the choice of sequential RSSI measurements over aggregate statistics. The observed pattern suggests that classifiers leverage both the starting signal strength and mid-trajectory measurements to infer movement direction, while later samples provide confirmatory information about the final position. This temporal dependency structure supports the use of sequence-based classification approaches for RSSI-based passenger detection.

H. Limitations and Considerations

Several limitations should be acknowledged. First, the controlled experimental environment, while designed to simulate public transport conditions, may not capture all sources of variability present in operational settings, such as passenger density fluctuations, vehicle movement, and diverse access point placements.

Second, the 10-second observation window, while suitable for capturing typical boarding and alighting actions, may be insufficient for detecting slower movements or hesitant passengers. Adaptive window lengths could potentially improve classification accuracy in such cases.

Third, the current approach assumes consistent device behaviour; however, variations in device hardware, operating system power management, and user-initiated WiFi state changes may affect RSSI reporting in practice.

Fourth, the limited sample size of the isolated dataset ($n = 159$) constrains the reliability of performance estimates for complex classifiers in that scenario. Future work should expand isolated data collection to enable more robust comparisons.

I. Practical Implications

The experimental findings hold meaningful implications for real-world deployment. The notable performance improvement observed in isolated conditions (MCC up to 0.907) suggests that signal interference is the chief limiting factor for classification accuracy. Deployment strategies that mitigate interference—such as dedicated frequency channels, directional antennas, or temporal multiplexing—could markedly enhance system performance.

Nevertheless, the classification performance achieved under noisy conditions (MCC > 0.77 with CatBoost) shows that RSSI-based movement classification remains viable as a complementary technology for passenger counting systems, even in challenging environments. The consistency of results across the combined and noisy-only scenarios lends confidence that models trained on realistic data will generalize well to operational settings.

VII. Contributions

This work puts forward the following contributions:

1) **Novel RSSI-based Movement Classification Framework:** A methodology exploiting temporal RSSI evolution to distinguish static and transitional movement patterns without precise localization.
2) **Purpose-Built Experimental Dataset:** Approximately 1,360 labelled samples across four movement classes, collected using four devices under isolated and noisy conditions.
3) **Comprehensive Classifier Evaluation:** Comparative analysis of 38 machine learning classifiers with multiple metrics and statistical validation across three random seeds.
4) **Feature Importance Analysis:** Identification that initial and mid-trajectory RSSI measurements contribute most to classification accuracy.
5) **Privacy-Preserving Approach:** Operation using only aggregate RSSI measurements without device identification or personal data.

Experimental results reaching over 81% accuracy underscore the practical viability of RSSI-based passenger movement sensing as a cost-effective complement to existing APC technologies.

VIII. Conclusions

This paper has presented a new approach for classifying passenger movements in public transport using temporal sequences of WiFi RSSI measurements. The proposed methodology enables non-intrusive distinction between four fundamental movement classes using only standard WiFi access point infrastructure.

The Gaussian Process classifier attained the highest performance, with an accuracy of 81.6%, recall of 81.6%, F1-score of 81.5%, and Matthews Correlation Coefficient of 0.756, validated across multiple random seeds. Support Vector Machines and regularized logistic regression variants yielded comparable results, confirming that both kernel-based and linear methods can effectively exploit the temporal structure of RSSI sequences.

Per-class analysis revealed that static states at the bus stop (BB) and transitional movements (AB, BA) are more readily distinguished owing to their characteristic signal patterns, whereas the static state inside the vehicle (AA) posed the greatest classification challenge because of its proximity-based similarity with the boarding class.

The results demonstrate that RSSI-based passenger movement classification offers a viable, cost-effective, and privacy-preserving complement to existing automatic passenger counting technologies, requiring no specialized hardware beyond standard WiFi infrastructure.

A. Future Work

Several directions for future research emerge from this work. First, validation in operational public transport environments is essential to gauge the impact of real-world factors such as vehicle movement, passenger density variations, and diverse access point configurations. Second, the integration of complementary sensor modalities such as accelerometer data or Bluetooth Low Energy beacons could enhance classification accuracy and robustness. Third, the development of adaptive observation windows that adjust to movement speed could improve detection of hesitant or slower passengers. Fourth, investigation of federated learning approaches would enable model improvement across multiple vehicles while preserving data privacy. Finally, the extension of the methodology to estimate complete origin-destination matrices through temporal aggregation of boarding and alighting events represents a natural progression toward comprehensive passenger flow analytics.

References

[1] Eurostat, "Urban and rural living in the eu," https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20230117-1, 2023, accessed: 2025-01-20.

[2] European Environment Agency, "Transport and environment report 2023," https://www.eea.europa.eu/publications/transport-and-environment-report-2023, 2023, accessed: 2025-01-20.

[3] B. Barabino, M. Di Francesco, and S. Mozzoni, "An offline framework for handling automatic passenger counting raw data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2443–2456, 2014.

[4] C. Pronello and X. R. Garzón Ruiz, "Evaluating the performance of video-based automated passenger counting systems in real-world conditions: A comparative study," *Sensors*, vol. 23, no. 18, p. 7719, 2023.

[5] S. Cerqueira, E. Arsenio, and R. Henriques, "Inference of dynamic origin-destination matrices with trip and transfer status from individual smart card data," *European Transport Research Review*, vol. 14, p. 49, 2022.

[6] M. Nitti, F. Pinna, L. Pintor, V. Pilloni, and B. Barabino, "iabacus: A wi-fi-based automatic bus passenger counting system," *Energies*, vol. 13, no. 6, p. 1446, 2020.

[7] A. Simoni, M. Mohori, and A. Hrovat, "Non-intrusive privacy-preserving approach for presence monitoring based on wifi probe requests," *Sensors*, vol. 23, no. 5, p. 2588, 2023.

[8] J. Guo, W. Zhuang, Y. Mao, and I. Ho, "Rssi-assisted csi-based passenger counting with multiple wi-fi receivers," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2025.

[9] C. Agualimpia-Arriaga, S. Govindasamy, B. Soni, C. Páez-Rueda, and A. Fajardo, "Rssi-based indoor localization using machine learning for wireless sensor networks: A recent review," in *IEEE ANDESCON 2024*, 2024.

[10] C. Wiboonsiriruk, E. Phaisangittisagul, C. Srisurangkul, and I. Kumazawa, "Efficient passenger counting in public transport based on machine learning," in *IEEE TENCON 2023*, 2023.

[11] T. A. Myrvoll, J. Håkegård, T. Matsui, and F. Septier, "Counting public transport passenger using wifi signatures of mobile devices," in *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2017.

[12] L. Fabre, C. Bayart, Y. Kone, O. Manout, and P. Bonnel, "A machine learning approach to estimate public transport ridership using wi-fi data," *IEEE Transactions on Intelligent Transportation Systems*, 2025.

[13] Y. Wang, H. Cheng, and M. Q.-H. Meng, "A learning-based sequence-to-sequence wifi fingerprinting framework for accurate pedestrian indoor localization using unconstrained rssi," *IEEE Internet of Things Journal*, 2025.

[14] V. Servizi, D. R. Persson, F. Pereira, H. Villadsen, P. Bækgaard, I. Peled, and O. A. Nielsen, ""is not the truth the truth?": Analyzing the impact of user validations for bus in/out detection in smartphone-based surveys," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[15] N. Singh, S. Choe, and R. Punmiya, "Machine learning based indoor localization using wi-fi rssi fingerprints: An overview," *IEEE Access*, vol. 9, pp. 127 150–127 174, 2021.

[16] B. Zholamanov, A. Saymbetov, M. Nurgaliyev, A. Bolatbek, G. Dosymbetova, N. Kuttybay, S. Orynbassar, A. Kapparova, N. Koshkarbay, and Ö. F. Beyca, "Rssi fingerprint-based indoor localization solutions using machine learning algorithms: A comprehensive review," *Smart Cities*, vol. 8, no. 5, p. 153, 2025.

[17] C. Jain, G. V. S. Sashank, V. N, and S. Markkandan, "Low-cost ble based indoor localization using rssi fingerprinting and machine learning," in *IEEE WiSPNET 2021*, 2021.

[18] D. Chicco and G. Jurman, "The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, 2020.

[19] ——, "The matthews correlation coefficient (mcc) should replace the roc auc as the standard metric for assessing binary classification," *BioData Mining*, vol. 16, no. 1, p. 4, 2023.