©2025 Published in 11th International Symposium on Intelligent Technologies in Engineering and Science 16-18 May 2025 (ISITES2025 Diyarbakır - Turkey)

Recent Advances in WiFi RSSI-Based Indoor Localization: A Machine Learning Perspective

[1]Sedat Çimen and *Sümeye Nur Karahan
[1]TT Mobil Istanbul, *[2]TT Mobil Ankara

Abstract

Indoor localization has become a critical component of modern smart environments, where traditional GPS technologies fail to provide reliable positioning. Among various techniques, Wi-Fi-based localization using Received Signal Strength Indicator (RSSI) has gained significant popularity due to its low cost and widespread infrastructure support. In recent years, machine learning (ML) algorithms have been increasingly applied to enhance the accuracy and adaptability of RSSI-based indoor positioning systems. This study presents a concise survey of research conducted between 2020 and 2025, focusing on supervised and deep learning models such as k-Nearest Neighbors, Support Vector Machines, Random Forests, and Convolutional Neural Networks. A comparative analysis of recent works is provided, highlighting their methodologies, datasets, and performance outcomes. Additionally, current challenges such as signal instability, device heterogeneity, and fingerprinting limitations are discussed. The paper concludes by outlining future research directions, including transfer learning and automated fingerprinting techniques.

Key words: Indoor localization, Wi-Fi RSSI, machine learning, fingerprinting

1. Introduction

In modern smart environments, especially with the growing number of Internet of Things (IoT) applications, accurate indoor positioning systems (IPS) have become increasingly important.
While Global Navigation Satellite Systems (GNSS) such as GPS provide reliable location services in outdoor environments, their performance drastically degrades indoors due to signal attenuation caused by obstacles like walls, ceilings, and dense structures [1].
As a result, indoor localization has emerged as a pivotal research field, enabling applications ranging from navigation assistance in shopping malls to asset tracking in healthcare facilities and smart factories [2]. Among various technologies proposed for indoor localization, Wi-Fi-based positioning has attracted considerable attention due to the ubiquity of Wi-Fi infrastructure in modern buildings and its cost-effectiveness [3]. One of the most widely used approaches in Wi-Fi-based localization is the utilization of Received Signal Strength Indicator (RSSI) values.
RSSI fingerprinting is favoured for its practicality and compatibility with existing infrastructure [4], where radio signal maps are created in an offline phase and matched with real-time measurements in an online phase.
Nevertheless, indoor environments pose several challenges for RSSI-based positioning, including signal fluctuations due to multipath propagation, device heterogeneity, and environmental dynamics [5].

*Corresponding author: Address: Faculty of Technology, Department of Civil Engineering Sakarya University of

Applied Sciences, 54187, Sakarya TURKEY. E-mail address: mehmets@subu.edu.tr, Phone: +0264 616 0000-01-02

To address these challenges and enhance localization accuracy, researchers have increasingly integrated Machine Learning (ML) techniques into RSSI-based systems.
These algorithms can effectively model the complex and non-linear relationship between RSSI measurements and physical locations, offering improvements over conventional statistical methods [6]. In recent years (2020–2025), there has been a significant increase in studies focused on RSSI-based indoor localization systems. These studies often employ supervised and deep learning methods, such as k-Nearest Neighbours (kNN), Support Vector Machines (SVM), Random Forests, and Convolutional Neural Networks (CNN), to enhance system performance and robustness [7][8]. However, a comprehensive and structured analysis of these developments remains limited. This paper provides a structured and focused review of recent advancements in Wi-Fi RSSI-based indoor localization, with particular emphasis on the role of machine learning algorithms.
We systematically review the methodologies, datasets, and performance outcomes reported in the latest literature. Furthermore, we highlight ongoing challenges and suggest promising future research directions to guide the continued development of robust indoor positioning systems.

The structure of the paper is as follows: Section 2 introduces the foundational concepts of RSSI-based indoor localization. Section 3 reviews the machine learning techniques employed in this context. Section 4 presents a comparative analysis of recent studies. Section 5 concludes the paper, while Section 6 discusses the main challenges and outlines future research directions.

2. Background and Fundamentals

Indoor positioning refers to the process of determining the location of objects or users within enclosed environments where satellite-based navigation systems such as GNSS are ineffective.
A variety of techniques have been developed to enable accurate indoor localization. These techniques include RSSI fingerprinting, Time of Arrival (ToA), Angle of Arrival (AoA), and Time Difference of Arrival (TDoA) [1]. Among these, Wi-Fi RSSI fingerprinting has become particularly popular due to its low cost, ease of implementation, and reliance on existing wireless infrastructure [8].

2.1. Received Signal Strength Indicator(RSSI)

RSSI is a measurement of the power present in a received radio signal, typically expressed in decibels relative to a milliwatt (dBm) [8]. In indoor localization systems, the RSSI value received from multiple Wi-Fi Access Points (APs) serves as a unique signature or "fingerprint" for a specific location. Since Wi-Fi signals are pervasive in most indoor environments, leveraging RSSI measurements provides a low-cost and infrastructure-free solution compared to alternative technologies such as Ultra Wide-Band (UWB) or Bluetooth Low Energy (BLE) [1].
However, RSSI measurements are susceptible to several sources of variability, including multipa th propagation, shadowing, and environmental changes (e.g., the presence of people, furniture reconfigurations) [1]. These variations can result in inaccurate positioning if not properly addressed during system design and implementation.

2.2. Wi-Fi Fingerprinting Approach

The Wi-Fi fingerprinting approach generally consists of two phases: an offline phase and an online phase. In the offline phase, RSSI values are collected at known reference points to build a radio map (fingerprint database) [7]. Each location is represented by a vector of RSSI measurements from surrounding APs. In the online phase, the user's device collects real-time RSSI values and attempts to match them to the closest fingerprint using machine learning or statistical algorithms [7]. Fingerprinting methods offer several advantages, including robustness to environmental noise and adaptability to complex indoor layouts. Despite these advantages, inherent limitations such as the necessity for labor-intensive site surveys and susceptibility to dynamic environmental conditions constrain their scalability [1].

3. Machine Learning Techniques in RSSI-Based Localization

Recent advances in machine learning (ML) have opened new avenues for enhancing Wi-Fi RSSI-based indoor localization, particularly in overcoming challenges posed by signal fluctuations, dynamic environments, and device heterogeneity. Traditional fingerprinting techniques often struggle to maintain accuracy under such conditions due to their limited capacity to model complex, non-linear relationships between RSSI vectors and physical locations. ML algorithms, on the other hand, offer a data-driven approach that improves localization accuracy and robustness by learning these intricate patterns [3]. Supervised learning models—including k-Nearest Neighbors (kNN), Support Vector Machines (SVM), Decision Trees, Random Forests, and Deep Neural Networks (DNN)—have been widely applied, demonstrating strong generalization performance even with noisy or incomplete data. Furthermore, ensemble and hybrid learning methods have been explored to increase the scalability and reliability of RSSI-based localization systems [4].

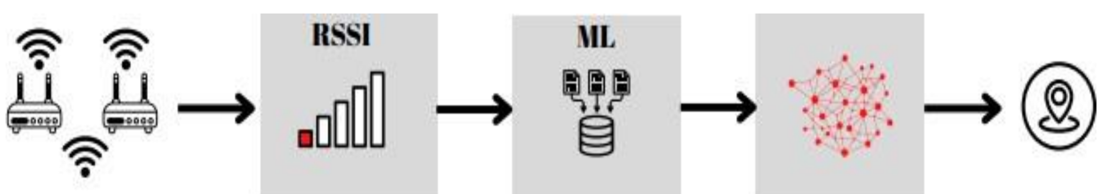

Figure 1. Workflow of Wi-Fi RSSI-based indoor localization using machine learning techniques

Wi-Fi signals received from multiple access points are converted into RSSI values, which are then used to build a fingerprint database. These RSSI vectors are processed by machine learning models to estimate the user's location accurately. The overall process is illustrated in Figure 1.

3.1. k-Nearest Neighbors

k-Nearest Neighbors (kNN) is a non-parametric, instance-based learning algorithm widely applied in RSSI fingerprinting [9]. In the context of indoor localization, kNN estimates the user's location by identifying the k closest fingerprints in the training dataset based on a predefined distance metric, typically Euclidean or Manhattan distance. **While kNN is easy to implement, it suffers from scalability issues and reduced performance in high-dimensional or sparse feature spaces [10].**
This limitation has been overcome through the development of enhanced versions of kNN, such as weighted kNN (WKNN), which have been shown to achieve optimal localization accuracy with minimal model complexity [11].

3.2. Support Vector Machines(SVM)

Support Vector Machines (SVM) are powerful supervised learning models that find an optimal hyperplane to separate data points from different classes [12]. In indoor localization, SVM has been employed both for classification tasks (e.g., predicting a room or floor) and for regression tasks (predicting precise coordinates) based on RSSI vectors.
SVM offers strong generalization performance, particularly when a limited amount of labeled data is available. Kernel-based SVM variants allow modeling of complex, non-linear relationships between RSSI values and physical locations [12]. However, SVMs can suffer from high computational demands during training, especially with large-scale fingerprint datasets.

3.3. Decision Trees and Random Forests

Decision Trees (DT) provide an interpretable model that recursively partitions the feature space based on feature thresholds . While individual decision trees are prone to overfitting, ensemble methods such as Random Forests (RF) have been shown to significantly improve localization robustness and accuracy [4].
Random Forests aggregate the predictions of multiple randomized decision trees, thereby reducing variance and improving generalization. In indoor positioning systems, RF models have demonstrated superior performance compared to simple classifiers like kNN, especially in noisy or dynamic environments [13].

3.4 Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN)

Deep learning methods, particularly Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN), have recently emerged as promising approaches for RSSI-based indoor localization [14]. DNNs can capture complex, hierarchical representations of RSSI data, enabling highly accurate localization even in challenging conditions.
CNNs, although originally developed for image processing tasks, have been adapted to treat RSSI fingerprints as structured inputs, allowing the networks to learn spatial correlations between access

points [15]. These models have been particularly effective when combined with data augmentation techniques to enhance model generalization.
However, the success of deep learning approaches heavily depends on the availability of large, diverse datasets, and they require significant computational resources for training and deployment.

3.5 Ensemble and Hybrid Methods

To further boost localization performance, researchers have explored ensemble and hybrid methods that combine multiple machine learning models [1]. For instance, ensemble strategies such as stacking or boosting leverage the strengths of different base learners to achieve higher accuracy and robustness.
Hybrid systems integrating conventional methods like kNN with advanced models like XGBoost or neural networks have shown considerable promise in adapting to dynamic indoor environments while maintaining computational efficiency [6] .

4. Discussion

Recent advances in RSSI-based indoor localization have been examined through a comparative evaluation of key studies. Table 1 highlights methodological trends, performance outcomes, and the evolution of machine learning approaches in this domain.

Table 1. Summary of Selected Recent Studies on RSSI-Based Indoor Localization (2020–2025)

No	Study	Year	Method	Dataset	Performance	Contribution
1	Kyounghyun Park et al. [6]	2023	Random Forest, XGBoost	Custom building (Samsung Galaxy S21+)	84.8%–100%	Handled missing data using one-hot encoding.
2	Jerry Landívar et al. [16]	2022	24 Regression Algorithms (Trilateration)	Custom Grid (400×200 cm²)	RMSE: 30–70 cm	Compared 24 different regression algorithms.
3	Kahraman Kostas et al. [5]	2022	XGBoost, ANN	13 public datasets	RMSE: 19.2	Proposed a venue-independent learning model.
4	Hera Neyaz et al. [7]	2024	SVM, KNN, Random Forest	UJIIndoorLoc	84%–99%	Designed separate classifiers for room and floor prediction.
5	M.W.P. Maduranga et al. [8]	2023	KNN, SVM, LDA	LoRa-based testbed	98%	Collected RSSI data from LoRa

						sensors compatible with 5G.
6	R.M.M.R. Rathnayake et al. [1]	2023	General ML Approaches	Literature review	-	Conducted a broad survey on RSSI-based localization for smart cities.
7	S. Hwang et al. [3]	2021	CNN-based	UJIIndoorLoc	91%	Created an RSSI heatmap and processed it using a CNN model.
8	Z. Chen et al. [4]	2023	CNN-LSTM hybrid	Indoor office environment	95%	Used time-series RSSI data to enhance location prediction.

Recent studies in RSSI-based indoor localization exhibit a clear shift toward ensemble learning and deep learning paradigms.
As evident from Table 1, algorithms such as Random Forest (RF) and XGBoost are repeatedly employed due to their robustness in handling noisy RSSI measurements and their generalizability across heterogeneous environments [5].
For instance, Kostas et al. demonstrate how ensemble models can maintain high accuracy even in unfamiliar venues by leveraging multi-environment training strategies.
Likewise, deep learning architectures—particularly CNN and CNN-LSTM hybrids—are effective in capturing spatial and temporal dependencies in RSSI fingerprints.
Z. Chen et al. [4] illustrate how incorporating time-series information improves the precision of indoor localization, achieving accuracies above 90% even in complex indoor layouts.
Another noteworthy trend involves dataset selection and diversity.
While standardized datasets such as UJIIndoorLoc continue to serve as useful benchmarks, recent efforts highlight the need for real-world, environment-specific datasets.
Studies like those by Maduranga et al. and Rathnayake et al. emphasize the importance of incorporating realistic conditions, including multipath interference, device variability, and architectural complexity, which better reflect practical deployment challenges.
Additionally, the comparative review points to the growing interest in generalizable and scalable models.
Most traditional models suffer significant accuracy drops when tested outside the training environment.
To overcome this, researchers have proposed venue-independent frameworks [5] and hybrid systems that combine lightweight models (e.g., kNN) with complex learners (e.g., DNN, XGBoost) [6].
These combinations aim to preserve accuracy while reducing computational cost, a critical factor

for mobile or embedded applications.
Despite these promising directions, several critical challenges remain unresolved:

- Multipath effects and environmental changes still degrade fingerprint reliability.
- Manual site surveys required for fingerprint construction continue to be labor-intensive and time-consuming.
- Device heterogeneity remains a major barrier for generalization [17].

Consequently, recent works increasingly emphasize the need for data augmentation, transfer learning, and automated fingerprinting strategies, which may serve as vital enablers for future indoor positioning systems.

5. Conclusion

This paper presented a structured survey of recent developments in Wi-Fi RSSI-based indoor localization systems powered by machine learning. By reviewing studies published between 2020 and 2025, we identified key trends such as the increasing use of ensemble learning, deep neural architectures, and hybrid modeling strategies. Our comparative analysis showed that while traditional approaches like kNN and SVM retain their relevance, modern techniques including CNNs, XGBoost, and CNN-LSTM hybrids provide notable improvements in localization accuracy, especially in complex or dynamic indoor environments. However, these advancements are faced with significant challenges, including signal variability, device heterogeneity, and the high cost of fingerprint data collection, which continue to impede large-scale and real-world deployments. Overcoming these limitations will require progress in areas such as transfer learning, automated fingerprinting, federated learning, and lightweight model design.
In the broader context, the development of models that are both generalizable and efficient in terms of resources, while also exhibiting a high degree of privacy awareness, is poised to play a pivotal role in facilitating the seamless integration of indoor positioning within smart building environments, IoT ecosystems, and pervasive computing infrastructures.
The potential of machine learning-driven indoor localization systems as central enablers of next-generation intelligent services is contingent on ongoing research and cross-disciplinary innovation.

6. Challenges and Future Research Directions

Indoor localization systems that rely on Wi-Fi RSSI have seen significant advancements, yet they continue to encounter persistent challenges that hinder their scalability, robustness, and practicality in real-world settings. One of the most critical issues is the inherent signal instability caused by multipath propagation, shadowing, and dynamic environmental conditions. These factors introduce significant noise into RSSI measurements, reducing localization precision, particularly in densely populated or frequently changing spaces. Device heterogeneity presents another major barrier, as RSSI readings from different smartphones or access points can vary substantially, undermining model generalizability across hardware platforms. Although calibration and normalization techniques have been proposed, achieving true device independence remains an unresolved issue. Additionally, the manual effort required for constructing fingerprint databases imposes a substantial burden on deployment. Traditional fingerprinting methods demand extensive on-site

surveys, which are time-consuming, labor-intensive, and impractical for large-scale or frequently updated environments. Environmental dynamics such as seasonal changes, renovations, or occupancy variations exacerbate this problem, requiring repeated data collection and model retraining. Furthermore, many machine learning models trained on specific buildings or layouts struggle to maintain accuracy when applied to unseen environments, highlighting the need for models with better generalization capabilities.
These ongoing challenges point to several promising directions for future research. Transfer learning and domain adaptation techniques have the potential to facilitate the reuse of models across different environments with minimal additional data, thereby reducing the necessity of repeated fingerprinting. Automated fingerprint generation, including the use of synthetic data via simulations or augmentation, offers a scalable solution to alleviate manual data collection. Federated learning frameworks may address device heterogeneity by enabling robust, privacy-preserving model training across multiple devices. Lastly, hybrid approaches that combine lightweight traditional algorithms with deep learning architectures can help balance interpretability, accuracy, and computational efficiency—making RSSI-based indoor localization systems more practical for deployment in real-world smart environments.

Acknowledgements

This work is supported by The Scientific and Technological Research Council of Türkiye (TÜBİTAK) 1515 Frontier R\&D Laboratories Support Program for Turk Telekom neXt Generation Technologies Lab (XGeNTT) under project number 5249902.

References

[1] Rathnayake RMMR, Maduranga MWP, Tilwari V, Dissanayake MB. RSSI and Machine Learning-Based Indoor Localization Systems for Smart Cities. Eng 2023;4(2):1468–94.

[2] Carlos Agualimpia-Arriaga, J. Rodríguez, F. Barrero. Machine Learning Based Indoor Localization Using Wireless Sensor Networks. IEEE Access 2023;11:18185–200.

[3] Hwang S, Lee J, Kang H. WiFi-based Indoor Positioning Using RSSI Heatmaps and CNN. Sensors 2021;21(22):8086.

[4] Chen Z, Zhang Y, He Y, Zhang S. Hybrid Deep Learning for Indoor Localization Using RSSI. Applied Sciences 2023;13(2):1104.

[5] Kostas K, et al. WiFi Based Distance Estimation Using Supervised Machine Learning. Proc IEEE Int Conf Indoor Positioning and Indoor Navigation (IPIN) 2022.

[6] Park K, et al. Improved RSSI-Based Indoor Localization Using Ensemble Learning. IEEE Access 2023.

[7] Neyaz H, Inamullah M, Beg MMS. Machine Learning Based Indoor Positioning System using Wi-Fi Fingerprinting Dataset. Proc IEEE ICECET 2024.

[8] Maduranga MWP, et al. Machine Learning-Based Indoor Localization System to Support 5G Location-Based Services. Proc 7th SLAICAI 2023.

[9] Hassoon SF, Kavian YS, Soorki MN. Indoor Localization using RSSI Based Supervised Machine Learning Approaches. *Proc IEEE ICACCCN* 2025.

[10] Machine Learning-Based Indoor Localization Using Wi-Fi RSSI Fingerprints: An Overview. *IEEE Access* 2021;9:127163–84. DOI: 10.1109/ACCESS.2021.3110430.

[11] Qian J, Komuro N, Kim W-S. Improving Position Estimation Accuracy by Reducing RSSI Offsets in BLE Fingerprinting-Based Indoor Positioning. *Proc IEEE ICMLC* 2022.

[12] Madhav, A., , Akhil, K. M. Rssi based localization and evaluation using support vector machine. In 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 509-515), 2022, September, IEEE.

[13] Hassen WF, Mezghani J. CNN-Based Approach for Indoor Positioning Services Using RSSI Fingerprinting. *Proc IEEE Conf on ITS* 2022;pp. 778–783.

[14] Jia B, Liu J, Feng T, Huang B, Baker T, Tawfik H. TTSL: An Indoor Localization Method Based on Temporal Convolutional Network Using Time-Series RSSI. *Comput Commun* 2022;193:293–301.

[15] Wu S, Huang W, Li M, Xu K. A Novel RSSI Fingerprint Positioning Method Based on Virtual AP and Convolutional Neural Network. *IEEE Sens J* 2022;22(7):6898–6909.

[16] Jerry Landívar et al., "Trilateration-Based Indoor Localization using Regression Methods," *Journal of Sensors*, 2022.

[17] Singh N, et al. Machine Learning Based Indoor Localization Using Wi-Fi RSSI Fingerprints: An Overview. *IEEE Access* 2021;9:127163–84.