Science Forum (Journal of Pure and Applied Sciences)

Journal Homepage: https://atbuscienceforum.com.ng/index.php/jpas

Wireless RSSI-Fingerprint and KNN Algorithm for improved Accuracy Indoor Positioning system

Abdoua Issa[1], Souley Boukari[1], Abdusalam Yau Gital[1], Aliyu Usman[2]

[1]Department of Computer Sciences, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria

[2]Federal College of Education (Tech), Gombe, Gombe State, Nigeria

Abstract

Indoor Positioning Systems (IPS) have attracted increasing attention due to the growing demand for location-based services in indoor environments where Global Positioning System (GPS) signals are often unreliable or unavailable. This study explores the application of the k-Nearest Neighbors (KNN) algorithm for indoor localization using Received Signal Strength Indicator (RSSI) values. The proposed framework leverages Wi-Fi signal strengths obtained from multiple Access Points (APs) to estimate the position of a target device within a predefined indoor area. A fingerprinting approach is adopted: during the offline phase, RSSI measurements are collected at known reference points to build a database, while in the online phase, the KNN algorithm matches current measurements with the *k* most similar signal patterns in the database and estimates the location based on the coordinates of these nearest neighbors. System performance is evaluated across different *k* values and varying environmental conditions, highlighting the trade-off between positioning accuracy and computational complexity. Results indicate that KNN provides a robust, low-cost, and computationally efficient solution well-suited for real-time indoor localization. Its simplicity and adaptability make it particularly promising for mobile and Internet of Things (IoT) applications. This research contributes to RSSI-based indoor positioning literature by assessing KNN performance in a 10.6 m × 5.8 m indoor space, using data from four Wi-Fi APs and 60 reference points. The findings reveal an average positioning error of approximately 0.5 meters, demonstrating the feasibility of the approach, while also acknowledging its limitations in handling signal variability and scalability in dynamic or larger environments.

Keywords

Fingerprinting Indoor Positioning System,
K-Nearest Neighbors (KNN),
Location Estimation,
Wi-Fi RSSI,

Corresponding Author: Abdoua Issa, Department of Computer Sciences, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria

Email: iabdoua.pg@atbu.edu.ng

I. INTRODUCTION

Indoor Positioning Systems (IPS) are increasingly important due to the growing need for location-based services in GPS-denied indoor environments. RSSI-based methods are widely adopted because they are simple, cost-effective, and compatible with existing wireless technologies like Wi-Fi and Bluetooth. However, these systems face challenges such as signal variability caused by environmental factors like interference, obstacles, and multipath effects, which reduce accuracy. Fingerprinting techniques, while effective, require extensive calibration and are difficult to scale in dynamic settings. Additionally, the lack of standardized evaluation frameworks and device-dependent signal readings complicates performance benchmarking and system deployment.

II. MATERIALS AND METHODS

Indoor positioning systems (IPS) are crucial for applications like indoor navigation, asset tracking, and context-aware services, especially where GPS signals are unavailable. RSSI-based techniques are among the most widely used alternatives due to their low cost and compatibility with existing Wi-Fi infrastructure (Mohsen et al., 2023). These methods typically use trilateration, fingerprinting, or path loss models, with fingerprinting offering high accuracy by comparing real-time data to a pre-built signal map (Jurdi et al., 2024).

However, fingerprinting requires extensive calibration and is sensitive to environmental changes, such as multipath effects, interference, and human movement. To improve adaptability and accuracy, machine learning algorithms like KNN (Lubis et al., 2020), SVM(Rathnayake et al., 2023), and deep neural networks (Mohsen et al., 2023) have been explored. While these models show promise, they often demand high computational resources and may need retraining when the environment changes. This study proposes a novel framework designed to overcome these limitations by improving robustness, scalability, and real-world applicability.

2.1 Methodology

This section outlines the models used in the current study, explaining their core principles and theoretical foundations. Each model is chosen for its ability to effectively process Wi-Fi signal data and deliver accurate indoor location estimates. The selection includes both traditional mathematical methods and modern machine learning algorithms, allowing for a comparison of different computational approaches. Key assumptions, strengths, and limitations of each model are examined in the context of indoor environments. The models are also evaluated based on their suitability for real-time applications, considering factors like accuracy(Lubis et al., 2020), efficiency, and ease of deployment. Overall, the chosen models provide a balanced and robust framework for developing a reliable indoor positioning system.

In indoor environments where GPS is ineffective, RSSI-based positioning offers a practical alternative by using Wi-Fi signal strength to estimate location. The K-Nearest Neighbors (KNN) algorithm is a popular choice for this task due to its simplicity and effectiveness. KNN is a non-parametric method that compares new RSSI readings with previously stored fingerprints and estimates position based on the average location of the K closest matches. It requires no model training and performs well in small to medium-sized areas with stable signal conditions. Its low computational cost and reliable accuracy make KNN a strong baseline in indoor positioning systems (Liu et al., 2020).

2.2 Knn algorithm

In indoor environments where GPS is unreliable or unavailable, alternative localization techniques are essential. One effective and straightforward approach is RSSI-based indoor positioning, which uses the Received Signal Strength Indicator (RSSI) from multiple Wi-Fi Access Points (APs) to estimate a device's location. A commonly used algorithm for this purpose is the K-Nearest Neighbors (KNN) algorithm. KNN is a non-parametric, instance-based learning method that does not require any model training. Instead, it relies on comparing new RSSI readings with previously collected data points (fingerprints) stored in a database. When a new RSSI reading is captured, KNN searches for the K most similar signal patterns in the database and calculates the estimated position by averaging the known coordinates of those K nearest neighbors.

This makes KNN particularly appealing for its simplicity, ease of implementation, and reasonable accuracy, especially in small to medium-sized indoor environments.(Masek et al., 2020) (Zhu et al., 2020). Due to its low computational cost and effectiveness in environments with relatively stable signal patterns, KNN remains a strong baseline method in the field of indoor positioning systems.

1. RSSI Indoor Positioning Using KNN – Step-by-Step

1. Environment Setup (Offline Phase)
   - **a.** Divide the indoor space into a grid or set of known reference points (RPs).
   - **b.** At each RP, collect multiple RSSI readings from nearby access points (Wi-Fi or BLE beacons).
   - **c.** Store the average RSSI values and their corresponding coordinates in a fingerprint database.
2. Fingerprint Database Creation
   - **a.** The database consists of entries like: [RSSI_1, RSSI_2, ..., RSSI_n] → (X, Y)

     Where RSSI values are the signal strengths from each transmitter, and (X, Y) are the coordinates of that location.
3. Online Phase: Real-Time Localization
   - **a.** The mobile device or target collects real-time RSSI readings from visible access points.
   - **b.** These values form the input vector for position estimation.
4. Distance Calculation
   - **a.** Compute the **Euclidean distance** (or another metric) between the real-time RSSI vector and each entry in the fingerprint database.
5. K-Nearest Neighbors Selection
   - **a.** Identify the **K** reference points (e.g., K = 3 or 5) with the smallest distances to the real-time RSSI vector.
6. Position Estimation
   - **a.** Estimate the device's position by averaging the coordinates of the K-nearest neighbors:

$$(X,Y) = \frac{1}{K}\sum_{i=1}^{i=k}(x_i, y_i)$$

one of the distances in (Dakkak et al., 2014) . these distances are in table n

7. Output

a. The estimated (X, Y) position is the final location prediction of the mobile device.

8. (Optional) Post-Processing

a. Apply smoothing techniques such as **Kalman Filter** to refine the location estimate over time.

2.3 Dataset and Data Collections

RSSI values are collected from multiple Wi-Fi access points at various known reference points within the indoor environment, typically using a mobile device or sensor(Simoes et al., 2020). Each data sample includes the RSSI readings from the access points along with the corresponding ground-truth coordinates (x, y), forming labeled training data. The dataset is then preprocessed by cleaning noisy data, normalizing RSSI values, and optionally splitting into training and testing sets to prepare it for input into machine learning models (Masek et al., 2020; Zhu et al., 2020).

The experimental evaluation was conducted in a controlled indoor environment measuring 10 meters in length and 5 meters in width. The test area resembled a typical office layout, containing desks, chairs, and other furniture that contributed to signal attenuation and multipath effects common in real-world indoor settings. To evaluate system robustness, several distinct scenarios were created, introducing variations such as: the presence or movement of individuals within the space, Altered orientation of the receiving device (e.g., facing north, south, etc.), Minor environmental changes (e.g., moved furniture, open/closed doors) (Ezzati Khatab et al., 2021).

These variations were designed to simulate the dynamic conditions typical of indoor environments. For the localization phase, the system employed a K-Nearest Neighbors (KNN) algorithm, which compared real-time RSSI vectors to the fingerprint database (Eleftherakis et al., 2024; Jurdi et al., 2024). The estimated position was computed as the weighted average of the k closest matches, based on RSSI similarity. These access points are strategically positioned to minimize blind spots and enhance localization accuracy. The exact spatial coordinates (x,y,z) as it is shown in *Table 1.* Each access point is pre-determined and recorded before beginning the experimentation process. This information serves as a reference for evaluating the performance of the positioning algorithm. The controlled placement and known positions of the access points contribute to the reliability and reproducibility of the test results.

Table 1: Access Points position

	x	y	z
AP1	4,00	10,00	1,00
AP2	1,00	10,00	1,00
AP3	4,00	0,00	1,00
AP4	0,00	1,00	0,00

The positions of the four access pointsAP1, AP2, AP3, and AP4 are listed in *Table 1*.

Table 2: Points of measurement

(PM)	(x;y)	(PM)	(x;y)
PM1	(5;10)	PM7	(5; 2)
PM2	(1; 10)	PM8	(3; 2)
PM3	(5; 1)	PM9	(3; 0)
PM4	(1; 1)	PM10	(1; 5)
PM5	(2; 5)	PM11	(1;3)
PM6	(5;5)		

In *Table 2,* Eleven measurement points are systematically designated within the infrastructure to serve as reference locations during the offline training phase. At each of these points, Received Signal Strength Indicator (RSSI) values are collected from all available access points. The measurements are performed under controlled conditions to ensure consistency and minimize environmental variability. These RSSI fingerprints are then stored in a database to be used later for position estimation during the online phase. The design and placement of these measurement points are critical for building an accurate and reliable fingerprinting model. The environmental test area is divided into fifty distinct reference points as given in *table 3*. These points are uniformly or strategically distributed throughout the environment to ensure comprehensive spatial coverage. At each reference point, signal measurements are collected to characterize the radio environment. This spatial granularity enhances the accuracy of the positioning system by providing a dense dataset for fingerprinting. The division into fifty (50) reference points allows for detailed analysis and validation of the localization algorithm's performance across different regions of the test environment.

Experimental set up

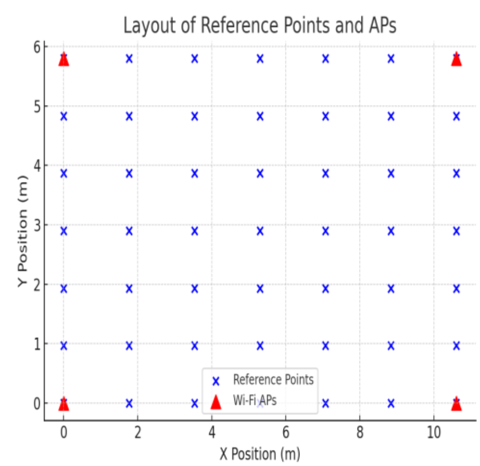

Figure 1 Test bed environment

Table 3: data collection scenario

III. RESULTS AND DISCUSSION

RP ID	X	Y	AP1	AP2	AP3	AP4
RP-01	0.0	0.0	-59	-70	-47	-63
RP-02	1.0	0.0	-48	-63	-73	-68
RP-03	2.0	0.0	-68	-43	-50	-50
…….	……..	…….	…….	…….	……..	……
RP-50	9.0	4.0	-46	-46	-43	-67

The k-Nearest Neighbors (KNN) algorithm showed reliable performance in estimating indoor positions based on Wi-Fi RSSI values. Despite its simplicity, KNN proved effective, though its accuracy can be affected by signal variability.

3.1 Results

Table 3: predicted positions

	Previous sys	P_err	Predicted	err
(5, 11)	(4, 11)	1.0	(4.40, 11)	0.60
(1, 10)	(2.4, 10.4)	1.45	(1.27, 10.2)	0.31
(5, 1)	(5,3.5)	1.58	(4.57, 1.1)	0.44
(1, 1)	(2.5, 1.7)	1.65	(0.93, 0.5)	0.80
(2, 5)	(1.7, 6.7)	1.44	(2.17, 2.6)	0.14
(5, 5)	(3.4, 3)	1.88	(4.7, 4.8)	0.30
(5, 2)	(3,4, 3)	1.55	(4.8, 1.8)	0.24
(3, 2)	(2, 3.2)	1.56	(2.7, 1.8)	0.34
(3, 10)	(2.4, 9)	1.16	(2.8, 9.8)	0.32
(1, 5)	(1.5, 4)	1.11	(1.8, 4.9)	0.21
(1, 3)	(1.3, 4)	1.4	(1, 2.8)	0.20

Table4 shows the estimated positions and errors produced by the K-Nearest Neighbors (KNN) algorithm at various indoor test points. Errors were calculated using the Euclidean distance between predicted and actual positions, providing a consistent metric for evaluation. Overall, KNN reduced positioning errors compared to the existing system, though a few points still showed notable deviations. The algorithm was evaluated with different *k* values (1 to 9), and the lowest errors occurred at *k* = 3, indicating an optimal balance between accuracy and robustness. Performance declined with higher *k* values due to the inclusion of distant, less relevant neighbors, confirming that moderate *k* values (3–5) are best suited for RSSI-based localization.

3.2 Discussion of Results

Table4 presents the predicted positions and corresponding errors obtained using the K-

Nearest Neighbors (KNN) algorithm across various test points in an indoor setting. For each measurement point, the positioning error was calculated using the Euclidean distance between the estimated and actual positions. The results indicate that the KNN algorithm consistently outperforms the existing system by significantly reducing the positioning error, in some cases by more than half. Notably, it showed high accuracy at specific points such as PM_2 and PM_5, achieving minimal errors of 0.31 m and 0.14 m, respectively. Despite these improvements, a few points revealed relatively larger errors, suggesting potential model limitations in certain spatial configurations. Further evaluation using different values of *k* demonstrated that *k = 3* yielded the lowest mean Euclidean error, striking a good balance between accuracy and resistance to noise. This outcome supports the preference for moderate *k* values in RSSI-based indoor localization, as they enhance precision while mitigating the effects of distant or noisy measurements. Its simplicity and effectiveness make it a reliable baseline, though its sensitivity to signal fluctuations can impact performance in dynamic or noisy settings.

REFERENCES

Eleftherakis, S., Santaromita, G., Rea, M., Costa-Perez, X., & Giustiniano, D. (2024). SPRING+: Smartphone Positioning from a Single WiFi Access Point. *IEEE Transactions on Mobile Computing*, *Section 3*, 1–18. https://doi.org/10.1109/TMC.2024.3367241

Ezzati Khatab, Z., Hajihoseini Gazestani, A., Ghorashi, S. A., & Ghavami, M. (2021). A fingerprint technique for indoor localization using autoencoder based semi-supervised deep extreme learning machine. *Signal Processing*, *181*(1). https://doi.org/10.1016/j.sigpro.2020.107915

Jurdi, R., Chen, H., Zhu, Y., Loong Ng, B., Dawar, N., Zhang, C., & Han, J. K. H. (2024). WhereArtThou: A WiFi-RTT-Based Indoor Positioning System. *IEEE Access*, *12*(February), 41084–41101. https://doi.org/10.1109/ACCESS.2024.3377237

Liu, R., Marakkalage, S. H., Padmal, M., Shaganan, T., Yuen, C., Guan, Y. L., & Tan, U. X. (2020). Collaborative SLAM Based on WiFi Fingerprint Similarity and Motion Information. *IEEE Internet of Things Journal*, *7*(3), 1826–1840. https://doi.org/10.1109/JIOT.2019.2957293

Lubis, Z., Sihombing, P., & Mawengkang, H. (2020). Optimization of K Value at the K-NN algorithm in clustering using the expectation maximization algorithm. *IOP Conference Series: Materials Science and Engineering*, *725*(1). https://doi.org/10.1088/1757-899X/725/1/012133

Masek, P., Sedlacek, P., Ometov, A., Mekyska, J., Mlynek, P., Hosek, J., & Komarov, M. (2020). Improving the Precision of Wireless Localization Algorithms: ML Techniques for Indoor Positioning. *2020 43rd International Conference on Telecommunications and Signal Processing, TSP 2020*, 589–594. https://doi.org/10.1109/TSP49548.2020.9163551

Mohsen, M., Rizk, H., & Youssef, M. (2023). Privacy-Preserving by Design: Indoor Positioning System Using Wi-Fi Passive

TDOA. *Proceedings - IEEE International Conference on Mobile Data Management*, *2023-July*, 221–230. https://doi.org/10.1109/MDM58254.2023.00045

Rathnayake, R. M. M. R., Maduranga, M. W. P., Tilwari, V., & Dissanayake, M. B. (2023). RSSI and Machine Learning-Based Indoor Localization Systems for Smart Cities. *Eng*, *4*(2), 1468–1494. https://doi.org/10.3390/eng4020085

Simoes, W. C. S. S., Silva, W. A. E., Lucena, M. M. De, Jazdi, N., & Lucena, V. F. De. (2020). A Hybrid Indoor Positioning System Using a Linear Weighted Policy Learner and Iterative PDR. *IEEE Access*, *8*, 43630–43656. https://doi.org/10.1109/ACCESS.2020.2977501

Zhu, X., Qu, W., Qiu, T., Zhao, L., Atiquzzaman, M., & Wu, D. O. (2020). Indoor Intelligent Fingerprint-Based Localization: Principles, Approaches and Challenges. *IEEE Communications Surveys and Tutorials*, *22*(4), 2634–2657.