IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 May 2024, accepted 4 June 2024, date of publication 11 June 2024, date of current version 19 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3412849

RESEARCH ARTICLE

Advance Path Loss Model for Distance Estimation Using LoRaWAN Network's Received Signal Strength Indicator (RSSI)

HOANG VO[1], VAN HOANG LONG NGUYEN[1], VAN LIC TRAN[2], FABIEN FERRERO[3], (Member, IEEE), FANG-YI LEE[4,5], AND MENG-HSUN TSAI[4,5]

[1]Faculty of Advance Science and Technology, The University of Danang–University of Science and Technology, Da Nang 550000, Vietnam
[2]Department of Electronic and Telecommunication Engineering, The University of Danang–University of Science and Technology, Da Nang 550000, Vietnam
[3]LEAT, CNRS, Universite Cote d'Azur, 06903 Sophia Antipolis, France
[4]Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 70101, Taiwan
[5]Department of Computer Science, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan

Corresponding authors: Van Lic Tran (tvlic@dut.udn.vn) and Meng-Hsun Tsai (tsaimh@cs.nycu.edu.tw)

This work was supported in part by the Science and Technology Development of The University of Da Nang under Project B2022-DN02-08. The work of Meng-Hsun Tsai was supported in part by the National Science and Technology Council (NSTC), Taiwan, under Grant 112-2221-E-A49-182; in part by the Industrial Technology Research Institute (ITRI); in part by the National Center for High-Performance Computing (NCHC); in part by the National Cheng Kung University (NCKU)-Accton Joint Research Center; and in part by the NCKU-Quanta Joint Research Center.

ABSTRACT This work introduces a novel approach to improve the precision of distance estimation in localization systems by using existing LoRaWAN and RSSI-based techniques. Despite the benefits of range and power efficiency, these systems exhibit limited accuracy in practical situations. To address the limitation, this study provides an innovative technique that greatly improves the precision of distance estimations, particularly in urban environments. The fundamental basis of this approach lies in the use of a dynamic path loss model. An additional element is to accommodate the varied and dynamic conditions of signal transmission in metropolitan areas. A better Kalman filter is also used in the study. This is important because it reduces the effects of multipath fading and environmental noise that often make RSSI-based localization in LoRaWAN networks less accurate. The study further examines the influence of the environmental exponent, also known as the path loss exponent, on the RSSI results and the precision of the distance measurements. This methodology achieves the average error under 1 meters for indoor environments and under 7 meters for outdoor environments. Finally, the Cumulative Density Function (CDF) shows 90 % of the distance estimation algorithm error for indoor environment is lower than 1.08 meters while for outdoor environment is lower than 7.55 meters. Based on these improvements, the introduced methodology not only enhances and improves existing approaches but also optimizes the precision and dependability of urban localization technologies, with substantial implications for a variety of practical applications.

INDEX TERMS Distance estimation, Kalman filter, localization, LoRaWAN, path loss.

I. INTRODUCTION

THE prediction of positioning is significant in several applications, including urban navigation, emergency response systems, and asset management in different industries [1], [2]. The primary obstacle in these applications is accurate distance measurement, especially in challenging circumstances such as indoor spaces and crowded urban regions. Although standard GPS technology, which is based on satellite signals, is reliable in outdoor environments [3], it is often inadequate in urban and interior areas due to signal obstacles [4]. The limitations of this technology, such as its accuracy margin of several meters, make it unsuitable for applications such as indoor navigation and asset tracking [5]. In addition, GPS devices are often power-consuming and costly when used for large-scale network deployments. Therefore, it is necessary

The associate editor coordinating the review of this manuscript and approving it for publication was Alessandro Pozzebon.

to investigate alternative ways in areas where GPS signals are ineffective [6].

Wireless sensor networks (WSNs) are now being used as a practical solution for indoor and urban localization [7]. These networks consist of distributed sensors that utilize several wireless technologies, including WiFi, Bluetooth, Radio Frequency Identification (RFID), Ultra Wide Band (UWB), and cellular networks. They are highly skilled at measuring distances and locations [8], [9]. Each technology has its own advantages and disadvantages, which make it appropriate for particular uses. With the different distance estimation measurements, including time of arrival (TOA), time difference of arrival (TDOA), and the received signal strength indicator (RSSI), each method serves a specific purpose. TOA (Time of Arrival) and TDOA (Time Difference of Arrival) are renowned for their exceptional accuracy; however, they require synchronized clocks [2], [8], [9], which in turn increase the complexity and costs of the system. However, RSSI is a simpler method, although it is less accurate and is susceptible to interference and environmental noise [10]. This study focuses primarily on improving the accuracy of the distance estimate of the LoRaWAN technology [11]. LoRaWAN stands out due to its extensive communication range and high-energy economy, making it particularly suitable for urban settings. LoRaWAN systems often use RSSI-based approaches because they can be easily adapted to various wireless communication protocols [9], [12].

This research presents a novel method to improve distance estimation accuracy in difficult indoor and urban environments, as mentioned, so we specifically focus on LoRaWAN and RSSI-based techniques. The essential aspect of this strategy involves creating a dynamic path loss model. This model is intricately crafted to effectively adapt to the changing conditions of signal propagation commonly found in urban environments. It responds dynamically to environmental factors such as urban clutter or building materials, which can greatly influence signal strength and quality [13], [14]. A new part is added to the model that makes it better adapted to different and changing urban signal settings. This makes distance estimates more accurate in situations where fixed models are not enough. There is noise in the environment and signal reflections from multiple paths that need to be dealt with in RSSI-based systems. This study combines the dynamic path loss model with an improved Kalman filter that does just that. The Kalman filter analyzes the incoming signal data, eliminating unwanted noise and rectifying distortions produced by multipath effects, thus yielding a more precise depiction of the true signal pathway. This enhancement leads to more accurate distance estimates. Furthermore, our study examines the relationship between the environmental exponent and changes in distances, which is crucial to understanding the mechanics of signal transmission in various environments. By gaining a comprehensive understanding of these dynamics, the suggested approach uses this information to enhance the process of calculating the distances with the creation of advanced algorithms that can intelligently modify the distance estimation model. These algorithms use predictive analysis of distance changes to continuously calibrate the device. This proactive adaptation ensures that the system maintains a high level of accuracy in diverse settings, successfully adapting to the various characteristics of the environment.

The main contributions of this paper are summarized as follows:

1) At the core of this approach is the development of a dynamic path loss model. This model is meticulously designed to respond to the fluctuating conditions of signal propagation typical in urban landscapes, adapting in real time to environmental factors such as urban clutter and building materials, which can significantly impact signal strength and quality. A new part is added to the model to make it more accurate at estimating distances in cities where static models are not good enough because signal conditions are always changing and being different.
2) Combined with the new element in the advance path loss model, the RSSI reference values at distance(d_0) have also been flexible adjusted. After a lot of testing and changes, it was found that dynamic changes in the RSSI reference value in the advanced path loss model algorithm make the distance estimate more accurate. In other studies, this change was not considered important.
3) Furthermore, research also investigates how the environmental exponent varies with changes in distances. This exploration is critical to understanding the dynamics of signal propagation in different settings. When a thorough understanding of these dynamics is complete, the proposed approach capitalizes on this knowledge to refine the distance calculation process.

The paper is organized as follows: Section II covers the technique, data collection, and algorithm design. Sections III and IV provide a comprehensive examination of the data and analysis, including the distribution of errors in various scenarios. Section V is the conclusion of our paper, which also proposes future exploration potentials on this topic.

II. RELATED WORK

Recent research has highlighted LoRaWAN's significant capabilities in localization, showing substantial advantages over traditional methods like GPS or Zigbee in terms of energy consumption and range while maintaining acceptable accuracy. These studies have primarily focused on Difference of Arrival (TDoA) and Received Signal Strength Indicator (RSSI) techniques for distance estimation, crucial for accurate localization [20].

One study explored a simulation combining TDoA and Angle of Arrival (AoA) techniques but lacked real-world testing [8]. Additionally, Spachos et al. noted that while RSSI values can be employed for indoor localization, they

TABLE 1. The summary results of some other related work.

Method	Approach	Test Area	Error avg. (m)
TDoA(outdoor)	-	0.65 x 0.65km	20–200 [23]
	-	1.74 x 1.74km	100 [24]
	Maximum Likehood Estimation (MLE)	10 x 10km	200 [25]
	Multilateral Dissection (MLD)	10 x 10km	500 [25]
RSSI(indoor)	Minimum Mean Square(MMS)	180 x 80m	20–30 [26]
RSSI(outdoor)	Weighted Least Squares (WLS)	150 x 100m	10-18 [21]
	-	120 x 200m	9–20 [27]
	Fingerprint algorithm	340 x 340m	24 [28]
	Fingerprint algorithm	7.28 x 7.28km	398 [29]
	Neural network	7.28 x 7.28km	357 [30]

are susceptible to significant errors due to environmental factors such as multipath effects and signal reflection, which can diminish the accuracy of the localization system [20]. Research by Kwasme and Ekin highlighted that indoor localization systems are challenged by dense multipath environments and non-line of sight conditions, which lead to increased localization errors. Their work suggests room for enhancements in both outdoor and indoor settings using RSSI [21]. They also demonstrated that the accuracy of TDoA geolocation could be enhanced through the application of the Kalman filter, although this was limited to devices in fixed positions [15]. The integration of TDoA and RSSI data using the k-nearest neighbors (kNN) algorithm alongside the more efficient LoRaWAN (V2) gateways, which achieved a median accuracy of 332.6 meters, shows the potential for mixed technology approaches. This aspect was explored further by Lam et al. in noisy outdoor environments, indicating practical improvements in field application [11]. Guo et al. investigated the integration of heterogeneous network technologies to improve location accuracy, emphasizing the importance of combining multiple data sources to enhance the reliability of distance estimations in variable urban environments [3]. Further foundational insights into the complexity of signal behavior in urban and indoor environments are provided by the studies of Han et al., which discussed new approaches to the geometry of TOA location [4], and Mani et al., who explored the influence of predicted and measured fingerprint on the accuracy of RSSI-based indoor location systems [5]. Further contributions by Raza et al. highlighted potential applications of LPWAN technologies, including LoRaWAN, for cost-effective alternatives to conventional GPS systems in urban areas [1]. Best practices in RSS measurements and ranging, critical for enhancing the performance of these technologies in dense urban settings, were also updated by these studies [2].

Table 1 presents a comparison of additional localization results concerning the mean error as documented in studies, as well as some details on the approach.

Our research builds upon these studies by proposing a novel approach that not only utilizes RSSI data but also enhances its accuracy through the development of a dynamic path loss model specifically tailored for urban environments. Unlike static models, our dynamic model adjusts in real-time to the fluctuating conditions of urban settings, effectively addressing the primary sources of error identified in previous studies. We also integrate a refined Kalman filter to mitigate the effects of environmental noise and multipath fading [15], which are prevalent in urban landscapes. This dual enhancement significantly reduces the average error in distance estimations, making our approach more robust and reliable for urban applications.

III. METHODOLOGY

To make the RSSI-based distance estimate better in LoRaWAN networks, a systematic four-step process is used. Figure 1 shows that the process includes setting up complex hardware, collecting dynamic data, analyzing the data in great detail, and making sure that the results are correct. The first phase sets up the LoRaWAN positioning system to operate within the frequency band AS923-2, using specific channels according to the collected data, ranging from 921.4 to 922.8 MHz for LoRa communication. Each channel is configured with a bandwidth of 125 kHz, in accordance with the regulation QCVN 122:2020/BTTTT.

The LoRa gateways, communicating with a processing unit via the Chirp Stack interface, are aligned with this configuration. In the second phase, a custom UCA educational board was set up to transmit signals and a laptop that serves as the data processor as describe in Figure 4. Regularly, the UCA board sends LoRa signals that the LoRa gateway receives and then forwards to the processing unit for data collection. Third, by collecting RSSI data systematically at different distances, it generates a comprehensive dataset that captures various signal strengths and environmental influences. This dataset is essential for precise modeling and analysis. Later, in the analytical phase, which is the fourth phase of our work, Python is utilized for data processing. The obtained

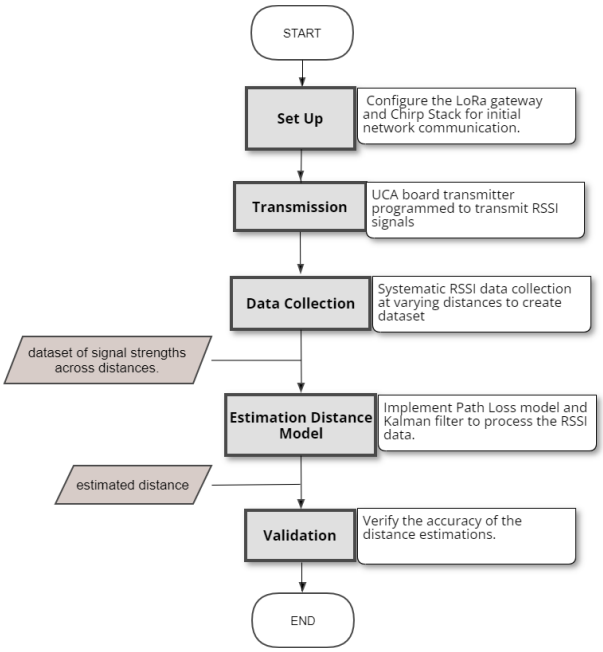


FIGURE 1. Flowchart of the methodology.

RSSI data is analyzed using an efficient Kalman filter and a path-loss model. This stage improves the precision of the signal and extracts the ambient exponent, a crucial factor for accurate distance estimations. In addition, it takes into account unpredictable fluctuations in RSSI values, which are represented as a Gaussian random variable. This is critical to the development of a reliable distance estimate model. This methodology guarantees flexibility and accuracy in addressing diverse urban signal situations, hence ensuring the model's correctness in different settings. The fourth and final part of our work checks how well the system works by validating it. This is done by estimating distances from sites that have not been visited before, processing RSSI signals using a better mathematical framework, and comparing the calculated distances with actual measurements to see how accurate the system is. Our methodology aims to improve the precision of urban and indoor localization technologies by integrating sophisticated hardware configuration, extensive data collection, meticulous analytical processing, and rigorous validation.

For the LoRa gateway device, Figure 2 shows the RAK7240 WisGate Edge Prime, a durable and adaptable device designed for secure IoT communication. It offers compatibility with LoRaWAN Version 1.0.2 and features dual LoRa concentrators for enhanced network density. Equipped with GPS, Wi-Fi, LTE, and Ethernet connectivity options, it ensures precise location tracking and versatile network integration. Experimentally, it operates within the 920-923 MHz band, adhering to regulatory standards.

Figure 3 presents the UCA Education Board, adapted for LoRaWAN applications. With compact dimensions and a miniaturized Printed Antenna, it allows frequency tuning from 845 to 950MHz, accommodating various global LoRaWAN bands.

FIGURE 2. LoRa gateway utilized in the setup.

FIGURE 3. UCA education board for signal generation.

A. DATA COLLECTION

The collection of Received Signal Strength Indicator (RSSI) data, which is crucial to improving distance estimation, was carried out in two different environments. The data gathering approach was consistent in both situations, beginning with the establishment of the LoRaWAN positioning system to ensure its proper functioning. Different RSSI signals, each associated with a distinct identifier, were periodically transmitted by the UCA Education Board. A methodical methodology was employed to collect RSSI data from different distances, obtaining 60 sets of RSSI signals at each specified distance to record oscillations in signal strength.

FIGURE 4. Data collecting setup.

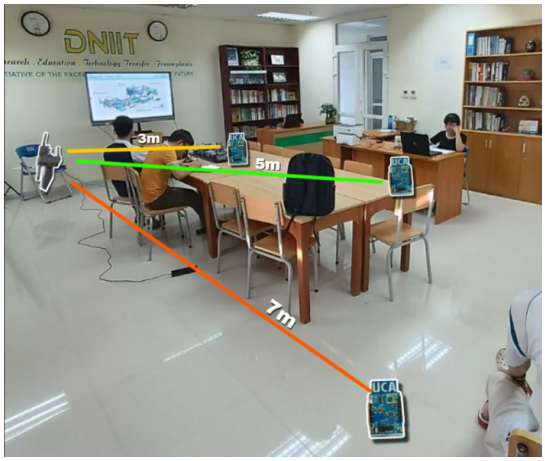


FIGURE 5. RSSI data collection in the indoor environment.

Data validation encompasses careful documentation of each distinct identification, associated distances, and pertinent metrics, which is then followed by a preprocessing phase to arrange the data for subsequent analysis. The meticulous gathering of data, conducted in various settings, both indoors and outdoors, has yielded a strong and indispensable data set that is crucial for developing an effective model for estimating distances. This methodology included the possibility of signal disruption, barriers, and fluctuations in signal intensity caused by varying distances within a given region.

1) INDOOR ENVIRONMENT

Data collection in the indoor environment occurred within a 10 m x 8 m room with some obstacles was shown in Figure 5, but the signal path remained largely in line of sight with minimal obstructions. Data were collected from three nodes at different distances of 3, 5, and 7 meters.

2) OUTDOOR ENVIRONMENT

In contrast, outdoor data collection was carried out in real-world conditions, incorporating obstacles such as trees and walls to replicate scenarios that could influence LoRa signal

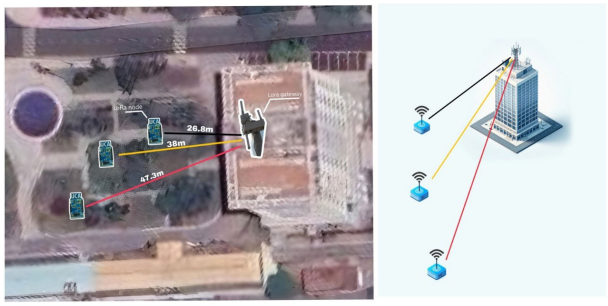


FIGURE 6. RSSI data collection in the outdoor environment.

propagation as was shown in Figure 6. Three nodes were used at different distances within a testing area of 150 m x 100 m, corresponding to 26.6, 38, and 47.3 meters.

B. DISTANCE ESTIMATION

In the distance estimation process, the RSSI values were analyzed and calculated using a simulation platform. An advanced path loss model is introduced based on the log-distance path loss model [19], as shown in the formula:

$$\mathrm{RSSI}(d) = \mathrm{RSSI}(d_0) - 10\eta \log_{10}\left(\frac{d}{d_0}\right) + T \qquad (1)$$

where η is the path loss exponent - a parameter that indicates how rapidly the signal strength decreases over distance in a specific environment [22], RSSI(d) is the value of RSSI at d distance from the transmitter, RSSI(d_0) is the value of RSSI at d_0 distance from the transmitter (which is often considered to be 1 meter in various studies [19], [22], and T is the Standard Deviation, a new element developed to enhance precision in distance estimation by minimizing errors generated by environmental factors such as reflection, fading, interference, noise, and shadowing. By incorporating T, the enhanced path loss model significantly improves its adaptability and accuracy in diverse conditions. Furthermore, the RSSI reference values at distance d_0 have been flexibly adjusted. Through extensive testing and modifications, it was found that the inclusion of T in the path loss model, a feature absent from other research, significantly reduces error. The equation is reorganized to compute the distance, symbolized by d. The path loss exponent, denoted as η, fluctuates based on the environmental conditions of the experiment.

$$d = 10^{\frac{\mathrm{RSSI}(d) - \mathrm{RSSI}(d_0) - T}{10}} \times d_0 \qquad (2)$$

C. PATH LOSS EXPONENT ESTIMATION

The Path Loss Exponent (PLE), also recognized as the path loss index or coefficient, is a key parameter in wireless communications. It describes how quickly the received signal strength (RSS) decreases as distance increases. PLE is essential to analyze and create distance estimation models for wireless systems, and its value varies according to the specific propagation environment. Generally, η (the symbol for PLE) ranges from 1.6 to 6. It is approximately 2 in free

space, between 2.7 and 3.5 in urban cellular areas, 3 to 5 in shadowed urban regions, 1.6 to 1.8 for in-building line-of-sight, 4 to 6 in obstructed in-building environments, and 2 to 3 in obstructed factory settings [22]. The equation of the propagation model is then rearranged to calculate the path loss exponent. The path loss exponent's value differs according to the environment used for the experiment.

$$\eta = \frac{\text{RSSI}(d_0) + T - \text{RSSI}(d)}{10 \log_{10}\left(\frac{d}{d_0}\right)} \tag{3}$$

D. KALMAN FILTER

The Kalman filter is a state estimator that provides an estimation of an unobserved variable using noise-containing measurements. It is a recursive algorithm that considers past measurements [18]. The task is to determine the actual RSSI using the measurements collected. The conventional Kalman filter operates under the assumption of linearity within its models. This implies that both the shift from the present state to the subsequent one and the translation of the state into measurement are expected to be linear transformations.

$$x_k = A_k x_{k-1} + \epsilon_k \tag{4}$$

The current state x_k is determined by combining the previous state x_{k-1} (using the state transition matrix A) and the noise ϵ_k. This noise is called process noise, as it is caused by the system itself. In the context of this RSSI filtering application, it is assumed that the device remains stationary. In addition, the position is considered static during the measurement period. In essence, a constant RSSI signal is expected throughout, with any deviations from this signal interpreted as noise. To reflect this assumption in the model, matrix A is defined as an identity matrix. This adjustment results in a significantly simplified model.

$$x_k = A_k x_{k-1} + \epsilon_k \approx x_{k-1} + \epsilon_k \tag{5}$$

The next step in incorporating the Kalman filter involves specifying the observation model, which describes the relationship between a specific state (x) and the corresponding measurement (z). The overall model can be expressed as follows.

$$z_k = C_k x_k + \delta_k \tag{6}$$

The transformation matrix C represents the transformation between the state and the measurements. The measurement noise, denoted as δ_k, is caused by faulty measurements. In our model, the received signal strength indicator (RSSI) is directly modeled, where the state and measurements match. It results in the subsequent simplified measurement model.

$$z_k = C_k x_k + \delta_k \approx x_k + \delta_k \tag{7}$$

By setting up the two transitions, the prediction step of the Kalman filter is established. This step describes the expected state of the system before any measurements are taken into account. With the presumption of a static system, the prediction step and the error covariance are defined as follows.

$$\hat{x}_k^- = \hat{x}_{k-1} \tag{8}$$

$$P_k^- = P_{K-1} + Q \tag{9}$$

The distinction between x and $\hat{x}$ lies in the fact that $\hat{x}$ represents our estimation, while x represents the true value of the state. The presence of a negative sign above $\hat{x}$ and P indicates that it is only a prediction which needs to incorporate additional information from the measurement. P determines the certainty of our estimate, which is based on the previous certainty and the process noise R that accounts for the noise generated by the system itself. The process noise R can be taken by calculating the variance of a set of RSSI values, which is the squared standard deviation. And the estimation value $\hat{x}$ is the mean RSSI of the data set. In the research, the process noise R is assigned a low value (e.g. 0.01) since it is claimed that most of the noise originates from the measurements. By utilizing the prediction estimate, Kalman gain is calculated. In the static system, a simplified version of the typical Kalman gain is employed.

$$K_k = P_k^-(P_k^- + R)^{-1} \tag{10}$$

serves as a factor that determines the balance between confidence in our estimation and reliability of the measurement, which is influenced by the presence of measurement noise (R) and estimation error (P). In situations where our prediction of the system is highly uncertain (i.e., P is large), it is advisable to place more trust in the measurement. Similarly, if the measurement noise is minimal (i.e. Q is low), it is also recommended to rely on the measurements. In such cases, the Kalman gain will be high. Conversely, if we have a high level of confidence in our prediction of the system and/or the measurement noise is substantial, the Kalman gain will be low. This directly affects the update step.

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - \hat{x}_k^-) \tag{11}$$

$$P_t = P_t^- - (K_k P_t^-) \tag{12}$$

During the update step, the final estimation of the system is calculated, denoted as $\hat{x}_k$, as well as the estimation uncertainty, P. The Kalman gain plays a crucial role in determining the extent to which the measurement is incorporated into the state estimate. A higher Kalman gain implies a greater integration of the measurement, while a lower Kalman gain indicates a higher level of trust in the estimation and a lesser reliance on the measurement information. This characteristic of the Kalman filter can be visualized as follows. At every step, the Kalman filter decides how much of the measurement it takes into account based on the certainty of the measurements.

In this research, Kalman filter's parameters are adjusted to fit the research objectives. Firstly, noise within the system is crucially examined through two parameters: Q, representing internal process noise, and R, reflecting measurement noise. Given the system's stability, Q is set low to signify minimal

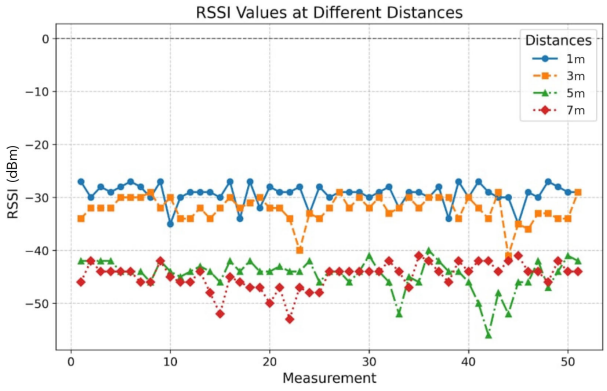


FIGURE 7. The RSSI values in the indoor environment.

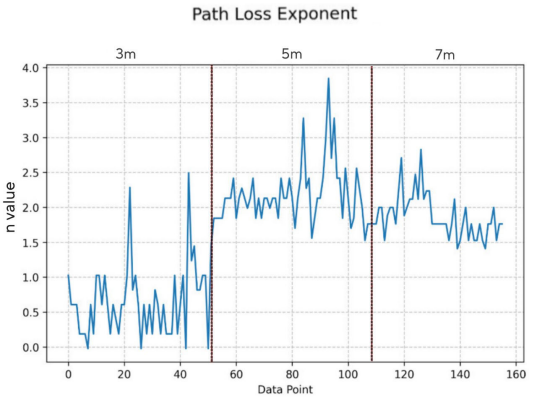


FIGURE 8. Measured environmental exponent.

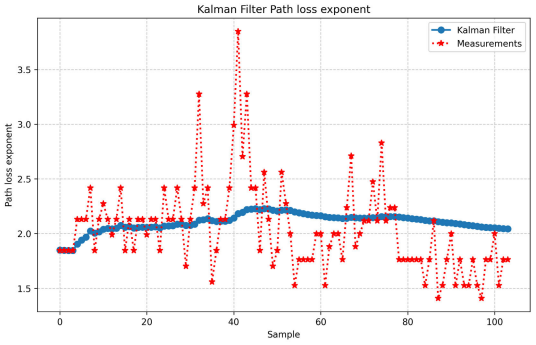


FIGURE 9. Filtered path loss exponents.

internal noise, while R is intentionally higher to account for measurement noise. For the analysis, Q is initialized at a low value, near zero, and R is estimated from the standard deviation of the Received Signal Strength Indicator (RSSI) signal. Secondly, for the RSSI filtering application, it is assumed that the device remains stationary. Furthermore, within the timeframe of the measurements, the gateway position is considered static. In essence, a constant RSSI signal is expected over time, with any variations being classified as noise. Thus, A and C are set to 1. Lastly, the initial estimation $\hat{x}_k^-$ is set to the average value of the RSSI signals in order to achieve a faster Kalman estimation.

E. ERROR ESTIMATION

The distance error is defined as the absolute of the distance between the estimated distance and the original one. The error is calculated as:

$$Error = |d_{est} - d_{real}| \tag{13}$$

In this equation, $d_{est}$is the estimated distance with T value and d_{real} is the actual distance from the transmitter to the receiver. Once the errors for all the tests performed are determined, an average can be taken. The resulting value of the mean error(ME) was then calculated as:

$$ME = \frac{1}{n}\sum_{i=1}^{n} Error_i \tag{14}$$

IV. RESULTS

A. INDOOR ENVIRONMENT

As Figure 7 shows, the farther the distance, the smaller the RSSI values. Also, the RSSI level varies a lot when we measure it. The RSSI fluctuation at each distance can reach up to 3 dBm. This is because signal transmission suffers from multipath fading. The RSSI values were measured in a small space that contained only tables and some devices that could not cause significant interference in the area, creating a low-noise environment for testing.

Then the Received Signal Strength Indicator (RSSI) in the log-distance model (formula 1) is utilized to estimate the path loss exponent η for the environment at distances d equal to 3, 5, and 7 meters, respectively, with a reference distance d_0 equal to 1 meter.

As shown in Figure 8, the path loss exponents are very small between 1 to 3 meters, indicating that the environment has a relatively weak impact on the Received Signal Strength Indicator (RSSI) value when the receiver is close to the transmitter. However, the η values increase significantly when measured at distances of 5 to 7 meters. This increase is due to factors such as signal reflection, fading, interference, and noise in the transmission environment.

Subsequently, Kalman Filter applied to estimate the path loss exponent of the environment. Since the path loss exponent has a weaker impact at close distances, for example, at 1 to 3 meters, the analysis focuses on the data collected at 5 to 7 meters. Figure 9 shows the comparison of the path loss exponent estimation with and without the Kalman filter. The red line represents the path loss exponent estimation without the Kalman filter, while the blue line represents the path loss exponent estimation with the Kalman filter.

Furthermore, the average of estimates provided by the Kalman filter is taken to determine the path loss exponent. The estimated path loss exponent of the indoor environment is **2.103**. For the next step, a test set of RSSI values from an unknown distance is collected, and then use the path loss exponent of the environment to estimate the real distance.

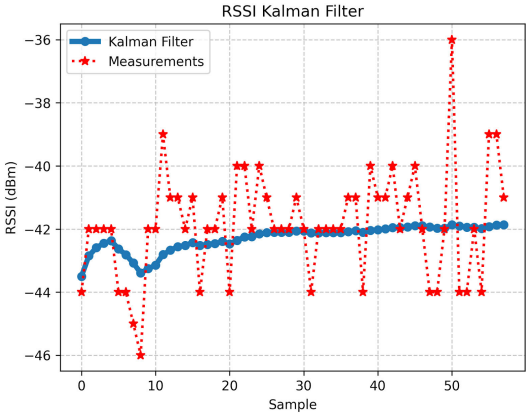


FIGURE 10. Filtered RSSI values of the first unknown distance.

TABLE 2. Distance estimation and error of indoor environment.

Raw data (dBm)	Estimated distance without T	Estimated distance with T	Real distance	Error
Average: -38.35 T value: 1.91	2.8 m	3.18 m	3.5 m	0.32
Average: -40.96 T value: 2.63	3.5 m	4.2 m	4 m	0.2
Average: -41.86 T value: 1.73	4.24 m	5.12 m	5 m	0.12
Average: -50.7 T value: 3.25	9.42 m	8.6 m	7.5 m	1.1

After we collect the new data, the Kalman filter is applied to minimize the fluctuation of the RSSI values. In figure 10, the red line is the RSSI values without the Kalman filter, whereas the blue line is the RSSI values with the Kalman filter.

As the figure 10 shows, the set of raw RSSI data varies from −39 dBm to −46 with a standard deviation of 1.73 dBm and an average value of −41.86 dBm. When the Kalman Filter is applied, the estimated RSSI at this unknown distance is −42.28 dBm with the standard deviation of only 0.5 dBm. For the next step, the distance estimate is made by formula 2. The result is **4.241 meters**. Furthermore, data are collected at a distance of **5 meters**. Later, the same process is applied to calculate the distance for 3 more different RSSI data sets. Resulted in the following table 2:

As Table 2 shows, the longer the distance from the receiver to the transmitter, the lower the RSSI value and the greater the overall standard deviation (T value). This can be easily explained by the increase of reflection, fading, interference, and noise from an environment when we increase the distance. While the estimations of 3.5, 4, and 5 meters take the reference value of 1 meters, the estimation of 7.5 meters actually takes the reference RSSI value of 5 meters. This is because when we look at the path loss exponent estimation in Figure 8, the value η of 5 to 7 meters is actually higher than that of 1 to 5 meters. Therefore, if an RSSI reference value of 1 meter was taken for all estimates, the precision would be reduced. Furthermore, the new T element in the path loss model also significantly decreases the system error.

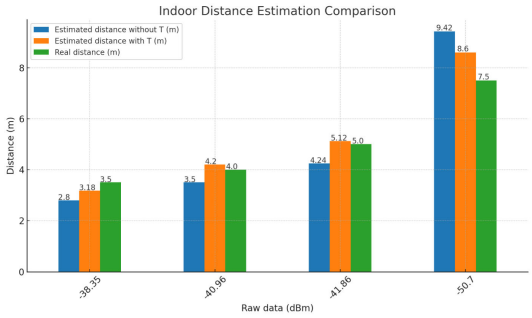


FIGURE 11. Distance estimation comparison graph of indoor environment.

Therefore, regardless of a highly fluctuated RSSI data set, a careful application of the path loss model as well as a well-programmed Kalman filter are very crucial in the calculation of distance.

Figure 11 presents a visual comparison of indoor distance estimations utilizing the proposed RSSI-based model with and without the T value. The inclusion of T significantly refines the distance estimations, as illustrated by the proximity of the 'Estimated distance with T' bars to the 'Real distance' bars. The raw RSSI data (in dBm) is used to estimate distances at various points, which are 3.5 meters, 4 meters, 5 meters, and 7.5 meters respectively. The graph clearly demonstrates the enhanced accuracy of our model with T. This marks a substantial improvement, showcasing the effectiveness of our advanced path loss model and Kalman filter in mitigating the impact of environmental noise and signal reflection within indoor settings.

Following that, we collected 50 data points at random positions within the testing range of 3 to 7 meters. The ME for indoor measurements is **0.565 meters**

B. OUTDOOR ENVIRONMENT

As Figure 10 illustrates, RSSI measurements collected outdoors at distances of 26.8 meters, 38 meters, and 47.3 meters show a reduction in signal strength as distance increases. The graph displays significant signal fluctuations, which is common in outdoor scenes due to obstacles such as walls, trees, and other structures that induce multipath fading. The signal of the outdoor environment suffers from noise more than the signal of the indoor signal in the previous experiment. Therefore, RSSI values in the outdoor environment fluctuate more significantly than in the indoor environment. In particular, the substantial variability at 38 meters suggests a dense presence of such obstacles, resulting in enhanced interference at this range and contributing to a challenging noise environment for testing and analysis. These data emphasize the need for environmental considerations in the effective management and mitigation of signal degradation and variability, which is crucial to maintaining strong and reliable LoRaWAN signal communication.

Then the RSSI in the log-distance model (formula 3) is applied to estimate the path loss exponent η for the

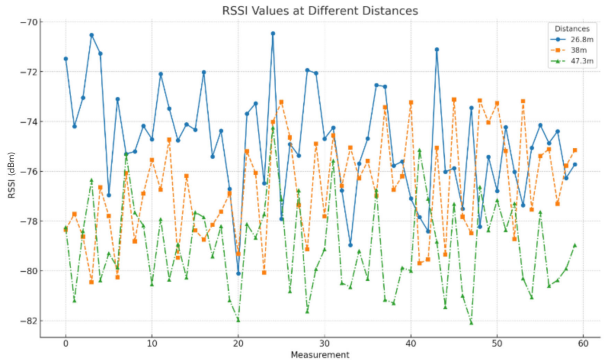


FIGURE 12. RSSI data collection in the outdoor environment.

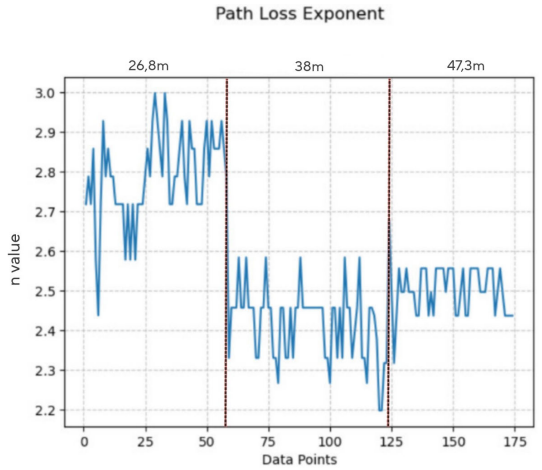


FIGURE 13. Path loss exponent measured of the outdoor environment.

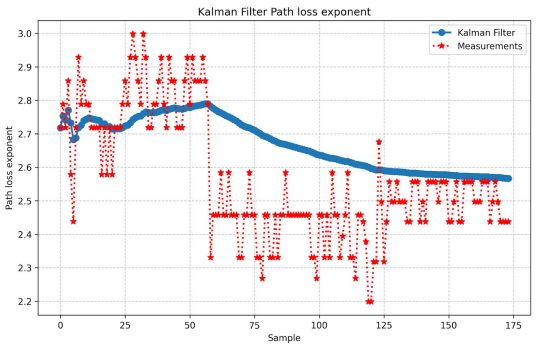


FIGURE 14. Filtered path loss exponent of the outdoor environment.

environment at distances d equal to 26.8, 38, 47.3 meters, respectively, and the reference distance d_0 equal to 1 meter.

As Figure 13 shows, the graph reveals data fluctuations due to environmental noise interference, requiring attention. Subsequently, the Kalman Filter is utilized to estimate the path loss exponent of the environment by isolating the true signal from the noise-induced variance. Figure 14 shows the comparison of the path loss exponent estimation graph with and without the Kalman filter.

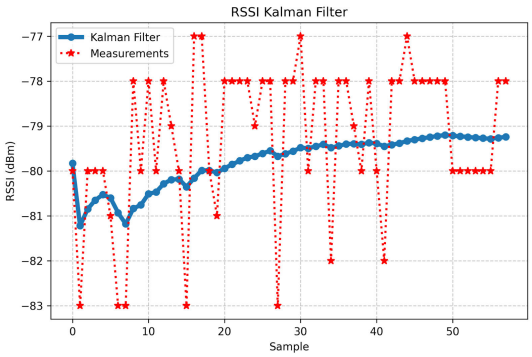


FIGURE 15. Filtered RSSI values of the unknown outdoor distance.

TABLE 3. Distance estimation and error of outdoor environment.

Raw data (dBm)	Estimated distance without T	Estimated distance with T	Real distance	Error
Average: -74.02 T value: 3.1	27.9 m	36.45 m	40 m	3.55
Average: -79.24 T value: 1.67	42.45 m	49.2 m	56.8 m	7.5
Average: -80.96 T value: 2.06	47.69 m	57 m	65.5 m	8.5
Average: -85.10 T value: 3.88	110.2 m	96.5 m	90 m	6.5

Next, the average of estimates provided by the Kalman filter is taken to determine the path loss exponent. The estimated path loss exponent for the outdoor environment is **2.691**. Subsequently, a test set of RSSI values is gathered from an unidentified distance and employ the path loss exponent of the environment to calculate the actual distance. Following data collection, Kalman Filter is utilized to reduce RSSI value fluctuations. Figure 15 illustrates this, the red line is the RSSI values without the Kalman filter whereas the blue line is the RSSI values with the Kalman Filter.

As depicted in the Figure 15, the initial RSSI dataset ranges from -77 to -83 dBm, showing a standard deviation of 2.15 dBm and an average value of -65.62 dBm. Following the application of the Kalman Filter, the estimated RSSI at this unidentified distance becomes -67.56 dBm, featuring a significantly reduced standard deviation of only 0.56 dBm. Subsequently, distance estimation is performed using the propagation model formula 2, which yields a result of **49.3 meters**. It is worth noting that the data was originally collected from a distance of **56.8 meters**, resulting in an error of 7.5 meters. Later, we used the same process to test 3 more RSSI data sets. The results are displayed in Table 3:

Figure 16 illustrates the comparative effectiveness of our distance estimation method in an outdoor environment, taking into account the dynamic path loss model with and without the adjustment value T. As observed, the estimations with the T value applied show a closer alignment with the real distances recorded, emphasizing the model's ability to contend with the complexities of outdoor signal propagation, including the effects of urban clutter and natural obstacles.

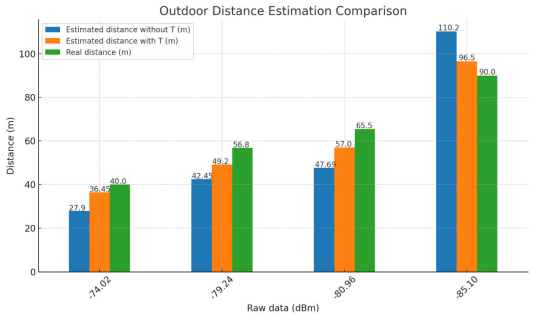


FIGURE 16. Distance estimation comparison graph of indoor environment.

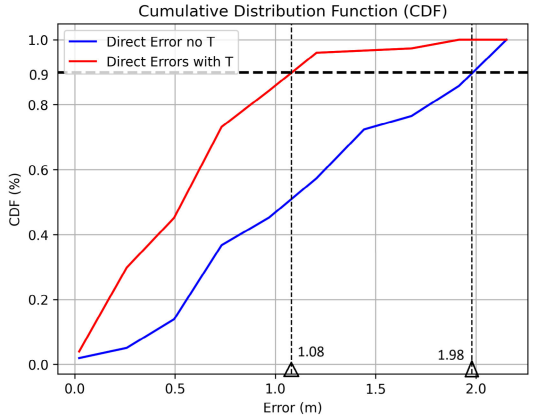


FIGURE 17. CDF distance error of indoor environment.

The raw RSSI data provided estimates for distances spanning 40 meters to 90 meters. The graph reveals our methodology's capacity to decrease the outdoor estimation error, with the largest improvement noted for distances beyond 50 meters, where environmental factors are most impactful.

Later, 50 points of data at random positions in the range of 30 to 90 meters were collected. The ME for outdoor measurement is **4.8 meters**.

C. CDF

Figure 17 and 18 shows the cumulative distribution of error for 50 random indoor and outdoor points, respectively. The measurement range for indoor is 3 to 7 meters, and for outdoor is 30 to 90 meters. The red line represents the cumulative distribution of error when the T value is utilized in the system and the blue line represents the cumulative distribution of error without the T value in the system.

As can be seen from Figure 17, the red line is higher than the blue line in the y axis. The proof is that 90% of the estimated distance has an error smaller than 1.08 meters for the system in which the T value is implemented, while for the system without the T value 90% of error is below 1.98 meters.

Figure 18 illustrates the cumulative error distribution in an outdoor setting. Similar to the earlier figure, the system incorporating the T value exhibits notably lower direct errors compared to the system lacking the T value. This is because,

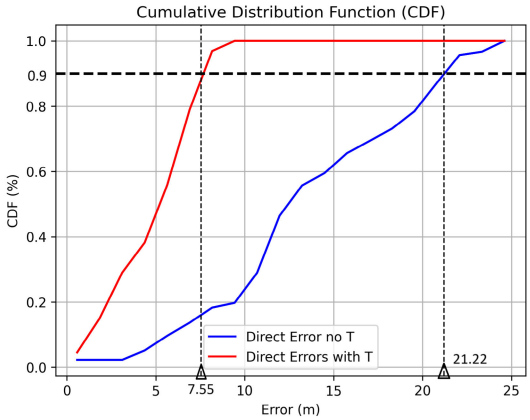


FIGURE 18. CDF distance error of outdoor environment.

in the system with the T value, 90% of the distance estimates have errors less than 7.55 meters, whereas in the system without the T value, 90% of the errors are under 21.22 meters.. Additionally, the maximum error recorded for the system incorporating the T value is 8.6 meters, in contrast to 24.6 meters for the system lacking the T value. This explains the presence of a long horizontal red line in the figure. It also suggests that the research method is more effective in noisy environments.

V. DISCUSSION

This paper presents findings that confirm the viability and illuminate the achievable outcomes and critical elements of distance estimation using RSSI-based methods in LoRaWAN networks. The Path Loss Exponent (PLE), indicating the rate at which received signal strength (RSS) diminishes with increasing distance, is significantly lower in indoor environments (2.103) compared to outdoor environments (2.691). This is because outdoor signals are more affected by noise than indoor signals. Additionally, the outdoor environment is cluttered with obstacles like walls, trees, and various structures that lead to multipath fading, whereas the indoor environment mainly consists of tables and a few devices, which do not significantly disrupt the signal.

Moreover, adding the new element T to the path loss model significantly enhances the system's accuracy. The indoor CDF figure shows that 90% of the distance estimates with the T value have an error smaller than 1.08 meters, compared to the system without the T value where 90% of the errors are less than 1.98 meters, nearly doubling the error rate of the system with the T value. Additionally, for outdoor CDF, 90% of the distance estimates in the system with the T value have errors below 7.55 meters, in contrast to the system lacking the T value, where 90% of the errors do not exceed 21.22 meters, almost tripling the error rate. These figures demonstrate that the addition of T to the path loss model substantially decreases the system's inaccuracies. This system is likely to provide more accurate distance measurements across various conditions,

particularly in outdoor environments, as it significantly reduces errors compared to indoor measurements.

Additionally, the Mean Error (ME) for indoor estimations stands at 0.565 meters, representing approximately 0.0807 to 0.188 of the measurement range of 3 to 7 meters. In contrast, the ME for outdoor estimations reaches 4.8 meters, which corresponds to 0.053 to 0.159 of the measurement range of 30 to 90 meters. These results indicate that the ME for outdoor measurements is significantly higher than for indoor measurements. Nevertheless, the proportion of error relative to the measurement range is marginally lower in outdoor measurements compared to indoor measurements.

Despite its positive outcomes, the system has several limitations. First, it necessitates a thorough examination of the environment, requiring a significant amount of data from the measurement area to gather all essential figures and conditions for the advanced path loss model and Kalman filter parameters to function correctly. Second, using the system to measure distances in locations outside the test area leads to greater inaccuracies, potentially causing system failures. Lastly, the system is only effective with stationary objects; its accuracy drops markedly when used with moving entities like vehicles or pedestrians. However, continuous research and development are being conducted to resolve these problems and improve the system's efficacy. In the future, efforts are being made to incorporate advanced machine learning and positioning techniques to improve the system efficiency

VI. CONCLUSION

A location estimate algorithm has been successfully created using the log-distance propagation model and Kalman filter, based on LoRa technology. The Received-Signal-Strength (RSS)-based technique successfully mitigates the accuracy limitations arising from environmental fluctuations, such as reflection, fading, interference, or noise. It has been determined that considering environmental fluctuations is equally important as accurately calculating distances. In indoor situations, the path loss exponent is at its lowest between lengths of 1 to 3 meters. However, as the coverage extends to distances of 5 and 7 meters, the influence of the surroundings on signal strength and stability becomes notably more pronounced, resulting in a substantial elevation of the path loss exponent. The exponent is determined by both the distance and the specific qualities of each environment, as demonstrated by the distinct path loss exponents in indoor (2.103) and outdoor (2.691) settings. Furthermore, the incorporation of the Kalman filter into the log-distance propagation model has resulted in improved precision and stability of the results. The utilized method resulted in ME (Mean Error) of 0.565 meters for the indoor estimation and 4.8 meters for the outside estimation. For the CDF at fixed position, 90% of the distance estimates have errors less than 1.08 meters for indoor environment and less than 7.55 meters for outdoor environment. Implementing the adjusted path loss model and the Kalman filter has demonstrated efficacy in reducing the influence of environmental factors on signal accuracy. In the future, the objective is to incorporate advanced machine learning and positioning techniques. such as fingerprinting or trilateration, to improve the precision of the system. The goal is to present a novel locating system that overcomes the limitations of GPS while maintaining exceptional precision and stability. This study represents progress toward achieving that goal, showcasing the potential of LoRa technology to improve future localization techniques.

REFERENCES

[1] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, 2nd Quart., 2017, doi: 10.1109/COMST.2017.2652320.

[2] A. Moradbeikie, A. Keshavarz, H. Rostami, S. Paiva, and S. Lopes, "Improvement of RSSI-based LoRaWAN localization using edge-AI," in *Science and Technologies for Smart Cities*. Cham, Switzerland: Springer, 2022, doi: 10.1007/978-3-031-06371-8_10.

[3] F. Lemic, A. Behboodi, J. Famaey, and R. Mathar, "Location-based discovery and vertical handover in heterogeneous low-power wide-area networks," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10150–10165, Dec. 2019, doi: 10.1109/JIOT.2019.2935804.

[4] X. Guo, Z. Chen, X. Hu, and X. Li, "Multi-source localization using time of arrival self-clustering method in wireless sensor networks," *IEEE Access*, vol. 7, pp. 82110–82121, 2019, doi: 10.1109/ACCESS.2019.2923771.

[5] R. S. Sinha and S.-H. Hwang, "Improved RSSI-based data augmentation technique for fingerprint indoor localisation," *Electronics*, vol. 9, no. 5, p. 851, May 2020, doi: 10.3390/electronics9050851.

[6] K. Chen, G. Tan, J. Cao, M. Lu, and X. Fan, "Modeling and improving the energy performance of GPS receivers for location services," *IEEE Sensors J.*, vol. 20, no. 8, pp. 4512–4523, Apr. 2020, doi: 10.1109/JSEN.2019.2962613.

[7] L. E. Marquez, A. Bahillo, M. Calle, and I. De Miguel, "Effects of body shadowing in LoRa localization systems," *IEEE Access*, vol. 11, pp. 9521–9528, 2023, doi: 10.1109/ACCESS.2023.3239896.

[8] R. Mani, A. Rios-Navarro, J.-L. Sevillano-Ramos, and N. Liouane, "Improved 3D localization algorithm for large scale wireless sensor networks," *Wireless Netw.*, pp. 1–16, Mar. 2023. [Online]. Available: https://link.springer.com/article/10.1007/s11276-023-03265-0, doi: 10.1007/s11276-023-03265-0.

[9] N. El-Sheimy and Y. Li, "Indoor navigation: State of the art and future trends," *Satell. Navigat.*, vol. 2, no. 1, p. 7, Dec. 2021, doi: 10.1186/s43020-021-00041-3.

[10] P. Savazzi, E. Goldoni, A. Vizziello, L. Favalli, and P. Gamba, "A Wiener-based RSSI localization algorithm exploiting modulation diversity in LoRa networks," *IEEE Sensors J.*, vol. 19, no. 24, pp. 12381–12388, Dec. 2019, doi: 10.1109/JSEN.2019.2936764.

[11] K.-H. Lam, C.-C. Cheung, and W.-C. Lee, "LoRa-based localization systems for noisy outdoor environment," in *Proc. IEEE 13th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Rome, Italy, Oct. 2017, pp. 278–284, doi: 10.1109/WIMOB.2017.8115843.

[12] T. Perković, L. D. Rodić, J. Sabić, and P. Solic, "Machine learning approach towards LoRaWAN indoor localization," *Electronics*, vol. 12, no. 2, p. 457, Jan. 2023, doi: 10.3390/electronics12020457.

[13] A. E. Ferreira, F. M. Ortiz, L. H. M. K. Costa, B. Foubert, I. Amadou, and N. Mitton, "A study of the LoRa signal propagation in forest, urban, and suburban environments," *Ann. Telecommun.*, vol. 75, nos. 7–8, pp. 333–351, Aug. 2020, doi: 10.1007/s12243-020-00789-w.

[14] M. Rademacher, H. Linka, T. Horstmann, and M. Henze, "Path loss in urban LoRa networks: A large-scale measurement study," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, Norman, OK, USA, Sep. 2021, pp. 1–6, doi: 10.1109/VTC2021-Fall52928.2021.9625531.

[15] C. H. Tseng and W.-J. Tsaur, "FFK: Fourier-transform fuzzy-c-means Kalman-filter based RSSI filtering mechanism for indoor positioning," *Sensors*, vol. 23, no. 19, p. 8274, Oct. 2023, doi: 10.3390/s23198274.

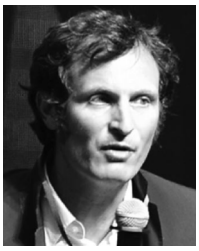
[16] R. Nogales, P. Tandazo, F. Mayorga, D. Guevara, and J. Vargas, "Open source cloud platform for academic systems monitoring software," in *Applied Technologies* (Communications in Computer and Information Science), vol. 1193, M. Botto-Tobar, M. Z. Vizuete, P. Torres-Carrión, S. M. León, G. P. Vásquez, and B. Durakovic, Eds. Cham, Switzerland: Springer, 2020, doi: 10.1007/978-3-030-42517-3_34.

[17] M. González-Palacio, D. Tobón-Vallejo, L. M. Sepulveda-Cano, S. Rua, G. Pau, and L. B. Le, "LoRaWAN path loss measurements in an urban scenario including environmental effects," *Data*, vol. 8, no. 1, p. 4, Dec. 2022, doi: 10.3390/data8010004.
[18] P. R. Gunjal, B. R. Gunjal, H. A. Shinde, S. M. Vanam, and S. S. Aher, "Moving object tracking using Kalman filter," in *Proc. Int. Conf. Adv. Commun. Comput. Technol. (ICACCT)*, Sangamner, India, Feb. 2018, pp. 544–547, doi: 10.1109/ICACCT.2018.8529402.
[19] N. A. M. Razali, M. H. Habaebi, N. F. Zulkurnain, M. R. Islam, and A. Zyoud, "The distribution of path loss exponent in 3D indoor environment," *Int. J. Appl. Eng. Res.*, vol. 12, no. 18, pp. 7154–7161, 2017.
[20] S. Sadowski and P. Spachos, "RSSI-based indoor localization with the Internet of Things," *IEEE Access*, vol. 6, pp. 30149–30161, 2018, doi: 10.1109/ACCESS.2018.2843325.
[21] H. Kwasme and S. Ekin, "RSSI-based localization using LoRaWAN technology," *IEEE Access*, vol. 7, pp. 99856–99866, 2019, doi: 10.1109/ACCESS.2019.2929212.
[22] V. Dharmadhikari, N. Pusalkar, and P. Ghare, "Path loss exponent estimation for wireless sensor node positioning: Practical approach," in *Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst. (ANTS)*, Indore, India, Dec. 2018, pp. 1–4, doi: 10.1109/ANTS.2018.8710123.
[23] LoRa Alliance. *LoRaWANGeolocation Whitepaper*. Accessed: Jun. 29, 2020. [Online]. Available: https://lora-alliance.org/sites/default/files/201804/geolocation_whitepaper.pdf
[24] B. C. Fargas and M. N. Petersen, "GPS-free geolocation using LoRa in low-power WANs," in *Proc. Global Internet Things Summit (GIoTS)*, Geneva, Switzerland, Jun. 2017, pp. 1–6, doi: 10.1109/GIOTS.2017.8016251.
[25] D. Bissett, "Analysing TDoA localisation in LoRa networks," M.S thesis, TU Delft Elect. Eng., Math. Comput. Sci., Delft Univ. Technology, Delft, The Netherlands, Oct. 2018.
[26] P. Manzoni, C. T. Calafate, J.-C. Cano, and E. Hernández-Orallo, "Indoor vehicles geolocalization using LoRaWAN," *Future Internet*, vol. 11, no. 6, p. 124, May 2019.
[27] A. Mackey and P. Spachos, "LoRa-based localization system for emergency services in GPS-less environments," in *Proc. IEEE INFOCOM Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Paris, France, Apr. 2019, pp. 939–944, doi: 10.1109/INFCOMW.2019.8845189.
[28] W. Choi, Y.-S. Chang, Y. Jung, and J. Song, "Low-power LoRa signal-based outdoor positioning using fingerprint algorithm," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 11, p. 440, Nov. 2018.
[29] M. Aernouts, R. Berkvens, K. Van Vlaenderen, and M. Weyn, "Sigfox and LoRaWAN datasets for fingerprint localization in large urban and rural areas," *Data*, vol. 3, no. 2, p. 13, Apr. 2018.
[30] G. G. Anagnostopoulos and A. Kalousis, "A reproducible comparison of RSSI fingerprinting localization methods using LoRaWAN," in *Proc. 16th Workshop Positioning, Navigat. Commun. (WPNC)*, Bremen, Germany, Oct. 2019, pp. 1–6.

HOANG VO was born in Da Nang, Vietnam, in 2004. He is currently pursuing the bachelor's degree with the Faculty of Advance Science and Technology, The University of Danang–University of Science and Technology, Vietnam.

VAN HOANG LONG NGUYEN was born in Vietnam, in 2004. He is currently pursuing the degree in embedded systems and IoT engineering with the Faculty of Advanced Science and Technology, The University of Danang–University of Science and Technology, Vietnam.

VAN LIC TRAN received the B.Eng. and M.Eng. degrees in electronics and telecommunications from The University of Danang–University of Science and Technology, Vietnam, in 2014 and 2019, respectively. He is currently pursuing the Ph.D. degree with the Universite Cote d'Azur, France. Since 2015, he has been a Lecturer with the Department of Electronics and Telecommunications, The University of Danang–University of Science and Technology. His main research interests include the Internet of Things, machine learning, LPWAN, LoRaWAN networks, and computer networking.



FABIEN FERRERO (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Nice-Sophia Antipolis, in 2007. From 2008 to 2009, he was with IMRA Europe (Aisin Seiki Research Center) as a Research Engineer and developed automotive antennas. In 2010, he was an Associate Professor with the Polytechnic School, Université Nice Sophia Antipolis. Since 2018, he has been a Full Professor with Université Côte d'Azur, where he is currently a member of the Laboratoire d'Electronique, Antennes et Telecommunications (LEAT). His current research interests include the design and measurement of millimetric antennas, the IoT systems, and space applications.

FANG-YI LEE received the B.Sc. degree in finance and international business from Fu Jen Catholic University, Taiwan, in 2016, and the M.Sc. degree from the Institution of Big Data, Taipei Medical University, Taiwan. She is currently pursuing the Ph.D. degree with the Institute of Computer Science and Engineering, National Yang Ming Chiao Tung University. In addition, she is gaining practical experience by working as an Intern at Accton Technology Company, recognized as one of Taiwan's leading networking firms. Her research interests include artificial intelligence, the Internet of Things, O-RAN, and the 6th generation mobile networks.

MENG-HSUN TSAI received the B.S., M.S., and Ph.D. degrees from National Chiao Tung University, in 2002, 2004, and 2009, respectively. He is currently a Professor with the Department of Computer Science, National Yang Ming Chiao Tung University. He was a Visiting Scholar with the University of Southern California, in 2012, for one month. His current research interests include mobile networks, software-defined networking, network security, and the Internet of Things. He was a recipient of the Exploration Research Award from the Pan Wen Yuan Foundation, in 2012, and the Outstanding Contribution Award from the IEEE Taipei Section, in 2010.