2022 9th International Conference on Digital Home (ICDH) | 978-1-6654-5478-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICDH57206.2022.00049

COOPERATIVE INDOOR LOCALIZATION SYSTEM BASED UWB AND RANDOM FOREST ALGORITHM IN COMPLICATED UNDERGROUND NLOS SCENARIO

Yaodong Yang[†],*Li Zhang*[*†], *Jingao Xu*[‡],*Danyang Li*[‡], *Jinhui Bao*[†], *Jieqing Tan*[†]

[†]School of Mathematics, Hefei University of Technology, Hefei,China
[‡]School of Software,Tsinghua University,Beijing,China
[*] Corresponding author e-mail:lizhang@hfut.edu.cn

ABSTRACT

In underground dangerous working environment, real-time positioning of person is required. For this application scenario that requires high positioning accuracy and anti-interference ability, ultra-wideband (UWB) positioning technology has more advantages than traditional positioning technology. However, underground scenarios are often complex and variable, and UWB measurements are affected by outliers in non-line of sight (NLOS) conditions, which can significantly reduce the performance of UWB localization systems. One cooperative indoor localization system based UWB and random forest algorithm has been proposed which innovatively applies random forest algorithm to Kalman filter's measurement update (UWB-RF). In particular, Taylor algorithm has been adopted to enhance estimation accuracy for the given localization system. We use indoor mobile localization experiments to compare the performance of the proposed approach with other NLOS mitigation methods. Experimental results show that the performance of the proposed UWB-RF approach is significantly better than the other methods.

***Index Terms*—** underground Indoor localization, NLOS mitigation, UWB, random forest

1. INTRODUCTION

Location information plays a pivotal role in modern society. Dangerous operation industries such as tunnels and mines require a positioning system with high positioning accuracy and high anti-interference ability to locate person in real time. When encountering a tunnel accident, the system can quickly find reliable location information of trapped people, to improve the emergency rescue capability and efficiency.

In the past decades, people have tried to apply various technologies to indoor localization, including WIFI, Bluetooth, IMU, sound, camera, and ultra-wideband (UWB). Among them, UWB is considered to be one of the most promising indoor positioning technologies because it contains many ideal characteristics like low energy consumption, centimeter-level range resolution, multi-path immunity, and certain obstacle penetration capability [1]. An indoor localization system based on UWB can obtain high localization accuracy under line-of-sight (LOS) environment, but it is easy interferenced by non-line of sight (NLOS) noise in a complex indoor environment, lead to poor positioning accuracy [2]. The identification and mitigation of NLOS have been a hot research topic in the field of UWB localization, many methods have been proposed to improve the localization accuracy [3].

The previous NLOS mitigation algorithm can be divided into two categories according to whether it can realize NLOS recognition or not. The first category is to ignore NLOS recognition and directly reduce the impact of NLOS on location estimation [4]. The second and most common approach is to achieve mitigation of NLOS in two steps. The first step is to identify NLOS and the second step is to eliminate the effect of NLOS on position estimation [3, 5, 6]. After NLOS has been identified, there are usually two strategies to achieve mitigation of NLOS. The first approach is to correct the distance error and use the revised distance estimate to locate [7]. The second strategy is to directly suppress the influence of NLOS on position estimation by specialized localization techniques, such as Taylor series least squares (TS-LS) method [8], Kalman filter (KF) [9], extended KF (EKF) [10], and cubature KF (CKF) [11].

In recent years, machine learning (ML) technology has also been studied and utilized as a new way of NLOS identification and mitigation. The main advantage is that they can be widely used without limited of prior knowledge. Previous investigations of NLOS recognition processes used different ML techniques to examine classifiers such as support vector machines (SVM) [12], Multilayer Perceptron (MLP) [13, 14], boosted decision tree (BDT) [14], and kernel principal component analysis [4], etc. In this paper, we propose cooperative indoor localization system based UWB and random forest (RF) algorithm, it is named name UWB-RF. The main contributions of this paper are summarized as follows:

1. This is the first time to incorporate RF algorithm and KF to reduce NLOS error and obtain more accurate distance estimation between the target node (TN) and corresponding anchor nodes (ANs).

978-1-6654-5478-0/22/$31.00 ©2022 IEEE
DOI 10.1109/ICDH57206.2022.00049

2. RF algorithm is used to analyze the underlying data of UWB to obtain accurate LOS and NLOS judgment.

3. Taylor algorithm has been applied to optimize filter's results and the superiority and adaptability of the proposed method have been verified by dynamic experiments in the evaluation scene. Experimental results show the localization accuracy has been significantly improved with high robustness and the strategy is effective.

The rest of this paper is organized as follows. In section 2, the system overview is given and related flow is described. In Section 3, the proposed UWB-RF position estimation algorithm and related knowledge are described in detail. Then, the performance of the proposed localization algorithm is evaluated in Section 4. Finally, some conclusions are given in Section 5.

2. SYSTEM OVERVIEW

Fig. 1 illustrates a functional overview of UWB-RF system. It is mainly divided into two modules: filter processing module and position calculation module. In the filter processing module, we need to collect some underlying information and labels of UWB offline and train the random forest model. Then, the UWB distance information is obtained through Time Difference of Arrival (TDOA), and the corresponding distance is determined in LOS or NLOS environment according to the underlying information and the trained random forest model. We use KF to achieve the mitigation of NLOS, and the distance information and environmental information are used to dynamically adjust during the measurement updating process. In the position calculation module, we use the least squares algorithm(LS) and Taylor algorithm to convert the distance information into position information, output the obtained position, update the distance between AN and TN, and use the updated distance for the next KF.

3. LOCATION CALCULATION

3.1. Random forest algorithm

In a complex indoor environment, UWB-ranging results obtained by TDOA are usually inaccurate under NLOS conditions. In an indoor localization system based on UWB, the distinction between the LOS and NLOS is very important. For now, ML approaches are an attractive solution to this problem.

We propose RF [15], a powerful integrated learning method based on classification and regression tree (CART), to distinguish LOS environment and NLOS environment. RF is a statistical learning theory, which uses the self-sampling method to extract multiple samples from the original samples, models the decision tree of each self-sampling, and then averages the final prediction results by combining the predictions of multiple decision trees. The model increases the diversity of decision trees by having a put-back sample and randomly changing the combination of predictors in different tree evolutions. It is used for classification and regression in many aspects and has good performance.

For NLOS identification, certain underlying parameters of UWB are required. The complete features extracted and preserved in our experiment are as follows:

1). the reported measured distance;
2). the amplitude of the first harmonic in the FP signal;
3). the maximum noise;
4). the standard noise;
5). the amplitude of the second harmonic in the FP signal;
6). the amplitude of the third harmonic in the FP signal;
7). the amplitude of the channel impulse response (CIR).

The modeling process of the RF algorithm is as follows: N estimators training sets are extracted from the original data set by bootstrap sampling technique. Each training set accounts for two-thirds of the original data set. During the training process, about one-third of each guide sample of the random forest is not mapped. This part of the data is called out-of-bag data. The next step is to create a regression tree for each guided training set. Construct N estimators' regression trees to form a "forest", but these regression trees are not pruned. RF algorithm increases the difference between regression models by constructing different training sets, to improving the extrapolation prediction ability of combined regression models. Through n-times of model training, we obtained a regression model sequence $\{t_1(x), t_2(x), \ldots, t_K(x)\}$, which is used to form a multiple regression model system (Forest). Then the prediction results of the regression tree of N estimators are collected and then we use a simple averaging strategy to calculate the value of the new sample. The final regression decision formula is as follows:

$$g_{rf}(x) = \frac{1}{K}\sum_{i=1}^{K} t_i(x), \tag{1}$$

where $g_{rf}(x)$ represents the combined regression model, $t_i(x)$ represents a single decision tree regression model, and K represents the number of regression trees (N estimators).

3.2. Least squares algorithm and Taylor algorithm

When we obtain a more accurate distance estimation, we can use LS [16] algorithm to estimate the position of TN. The available observation equation of the UWB localization system composed of n ANs is modeled as follows:

$$\begin{cases} \sqrt{\left(x_1^{AN}-x\right)^2+\left(y_1^{AN}-y\right)^2}=\widetilde{d}_1 \\ \sqrt{\left(x_2^{AN}-x\right)^2+\left(y_2^{AN}-y\right)^2}=\widetilde{d}_2 \\ \cdots \\ \sqrt{\left(x_n^{AN}-x\right)^2+\left(y_n^{AN}-y\right)^2}=\widetilde{d}_n, \end{cases} \tag{2}$$

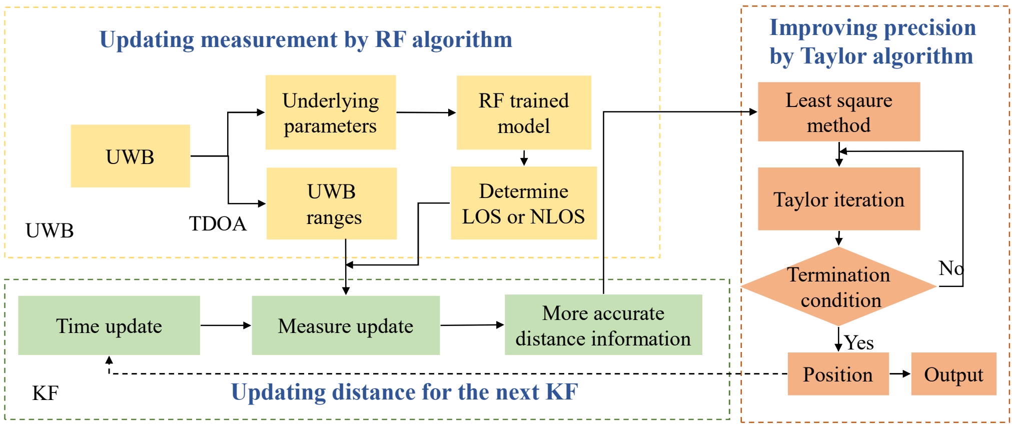

Fig. 1. System Overview

where $\widetilde{d}_1$,$\widetilde{d}_2$,..., $\widetilde{d}_n$ represent the distance between the calibrated TN and the corresponding AN, respectively. $\left(x_i^{AN}, y_i^{AN}\right)$ represent the coordinates of the ith AN through $i = 1, 2, \ldots, n$.

Conversion of (2) to matrix form produces:

$$AX = D, \tag{3}$$

The position of TN can be obtained by LS:

$$X = \left(A^T A\right)^{-1} A^T D. \tag{4}$$

Taylor algorithm is a recursive algorithm, for the initial position obtained by LS algorithm, it gradually converges to the estimated position through the recursive method, and gets more accurate results [17]. For each set of measured values, first the Taylor algorithm performs Taylor expansion on target position and then ignores the second-order or higher part to carry out the recursive calculation. The main idea of the algorithm is to revise the estimated value of the undetermined tag position through continuous iteration, and finally approach the real tag position coordinates gradually. The calculation process can be described as follows:

Assume that the real coordinate of TN is (x, y), and the obtained by iteration is (x_e, y_e). The relationship between the real coordinate and the estimated coordinate is shown in Equation (5).

$$\begin{cases} x = x_e + \Delta x \\ y = y_e + \Delta y \end{cases} \tag{5}$$

The algorithm obtains the initial coordinate (x_0, y_0) through LS for iterative operation. Taylor expansion is performed in Equation (2), and the second-order or above part are ignored:

$$\widetilde{d}_i = \widetilde{d}_i\Big|_{x,y} + \frac{\partial \widetilde{d}_i}{\partial x}\bigg|_{x,y} \Delta x + \frac{\partial \widetilde{d}_i}{\partial y}\bigg|_{x,y} \Delta y + \varepsilon_i. \tag{6}$$

the error of the measured value is expressed as:

$$\varphi = h_t - G_t \Delta. \tag{7}$$

Let $\varphi = 0$, then according to the LS algorithm we can get:

$$\Delta = \left(G_t^T G_t\right)^{-1} G_t^T h_t. \tag{8}$$

In the next iteration, set $x_0 = x_0 + \Delta x$, $y_0 = y_0 + \Delta y$ until $|\Delta x|^2 + |\Delta y|^2 < P$ is satisfied and P is the given threshold, then the operation result at this time is the coordinate value of TN obtained.

3.3. Kalman filter

KF can estimate the state vector of the moving target using the observation distance data [18], and then the real distance value can be estimated. Suppose that the state vector of a mobile device can be represented by Equation (9).

$$X(k+1) = \Phi X(k) + W(k) \tag{9}$$

Where $X(k) = \left[d_1(k), \ldots, d_n(k), \dot{d}_1(k), \ldots, \dot{d}_n(k)\right]^T$ represents the state vector of TN at time k ,$W(k)$ is the driving noise vector with covariance matrix $Q = \sigma_W^2$, and

$$\Phi = \begin{bmatrix} I & \Delta t * I \\ 0 * I & I \end{bmatrix}, \tag{10}$$

I is the identity matrix of $n * n$.

The measurement process is shown in the formula below

$$Z(k) = HX(k) + U(k), \tag{11}$$

where $Z(k) = \left[\tilde{d}_1(k), \tilde{d}_2(k), \ldots, \tilde{d}_n(k)\right]$ is the measurement vector, $U(k)$ is the measurement noise with covariance $R = \sigma_u^2$, and

$$H = \begin{bmatrix} I \\ 0 * I \end{bmatrix}. \tag{12}$$

When we perform KF, we extract the underlying parameters and determine whether each distance is LOS or NLOS through RF algorithm, so as to dynamically adjust $R(k)$. We can conduct experiments in the local environment in advance to obtain r_{los} and r_{nlos} , then select the corresponding value in the corresponding diagonal part of $R(k)$. And when an original distance data is detected in NLOS, because the mean error of NLOS is positive, we subtract an initial distance error δ_d from the initial distance, so as to achieve the effect of improving the accuracy.

When we update distance information each time through iterative operation, we can use Taylor algorithm to get the position of TP at this moment, and we can calculate the distance between TP and each AP $d_1(k), d_2(k), \ldots, d_n(k)$, take this distance as the next iteration of KF. The whole UWB-RF process is shown below.

UWB-RF approach

Input: Set $x_{0|0}, P_{0|0}, Q, \mathbf{r}_{\text{los}}, \mathbf{r}_{\text{nlos}}, \delta_d$
Time update
$\hat{X}(k \mid k-1) = \Phi \hat{X}(k-1 \mid k-1)$
$P(k \mid k-1) = \Phi P(k-1 \mid k-1) \Phi^T + Q$
Measure update
The LOS or NLOS information is obtained by random forest model
for i from 1 to n :
if $\tilde{d}_i(k)$ in LOS:
$R_{i,i} = r_{los}$
if $\tilde{d}_i(k)$ in NLOS:
$R_{i,i} = r_{nlos}$
$\tilde{d}_i(k) = \tilde{d}_i(k) - \delta d$
$K(k) = P(k \mid k-1) H^T \left[H P(k \mid k-1) H^T + R(k) \right]^{-1}$
$\hat{X}(k \mid k) = \hat{X}(k \mid k-1) + K(\mathrm{k})[Z(\mathrm{k}) - H\hat{X}(k \mid k-1)]$
$P(k \mid k) = P(k \mid k-1) - K(\mathrm{k}) H P(k \mid k-1)$
End for
$(x_k, y_k) =$ Taylor Algorithm$(d_1(k), d_2(k), \ldots, d_n(k))$
Calculate $\left(\hat{d}_1(k), \hat{d}_2(k), \ldots, \hat{d}_n(k)\right)$ by (x_k, y_k) and $\left(x_i^{AN}, y_i^{AN}\right), i = 1, 2, \ldots, n.$
Update $\hat{X}(k \mid k)$ through $\left(\hat{d}_1(k), \hat{d}_2(k), \ldots, \hat{d}_n(k)\right)$ by (x_k, y_k)
Output: (x_k, y_k)

---

4. EXPERIMENTAL SETUP AND RESULTS

4.1. Experimental setup

DWM1000 module can be used in TWR or TDOA localization systems to locate targets with an accuracy less than 10 cm. The module supports data transfer rates of up to 6.8 Mbps. DWM1000 modules are selected as TN and ANs for their high performance and reliability. In order to obtain the location information of the intelligent robot, we install TN on the intelligent robot body and four ANS on fixed positions in the room.

In order to verify the effectiveness, adaptability and robustness of our proposed method, we conduct an experiment on the one of the floors of a building in Hefei University of Technology. Four ANs are configured as rectangular deployment mode and distributed in the four corners of the room. In the dynamic experiment, TN moves along a preset trajectory. Respectively fixed on the tripod of position coordinates AN1 (0.0, 0.0), AN2 (0.0, 4.8), AN3 (6.4, 4.8) and AN4 (6.4, 0.0). The heights of AN and TN are fixed, and the distance information we get is obtained by modifying the UWB ranging information, as shown in Equation (13).

$$\tilde{d}_i = \sqrt{\bar{d}_i^2 - (z_i - z_{TN})^2} \tag{13}$$

Where $\bar{d}_i$, z_i and z_{TN} are the distances measured from TN of the ith AN and the heights of the ith AN and TN respectively.

4.2. Results analysis

In order to highlight the superiority of this technology, we adopt widely used root mean square error (RMSE) and absolute error and as the main indexes to evaluate the localization performance.

In order to demonstrate the rationality and adaptability of the proposed UWB-RF scheme, we also optimize the results of LS and total least squares (TLS) [19] algorithms. The algorithm is tested with the data before and after KF to further fully demonstrate the advantages of the proposed UWB-RF technique. In addition, we use "RF" to indicate that RF algorithm is added to KF process for NLOS identification. The symbols of the methods we use are as follows: UWB-RF, N-Taylor, RF-TLS, RF-LS and Taylor, TLS, LS. Here N-Taylor indicates that KF is used but RF algorithm is not used.

The comparison of TN position estimation errors generated by different localization methods in the evaluation experiment is shown in Fig. 2. And the cumulative distribution function (CDF) of localization error is shown in Fig. 3. It can be seen intuitively that UWB-RF method has the smallest localization error, while the localization error of TLS is the biggest. The proposed UWB-RF can effectively improve localization accuracy. Another observation can be obtained from these results: compared with RF-TLS and RF-LS, UWB-RF has smaller localization error, indicating that Taylor algorithm can further efficiently and effectively improve localization accuracy. It is worth noting that UWB-RF method not only has much smaller localization error than other methods, but also has very gentle localization error fluctuation range, which highlights its superior localization performance and robustness, makes it more suitable for underground applications.

In order to further investigate the localization performance achieved by the proposed UWB-RF method and other

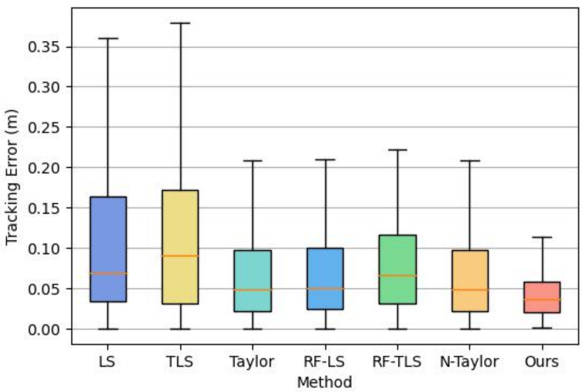

Fig. 2. Box plots of localization errors from different methods

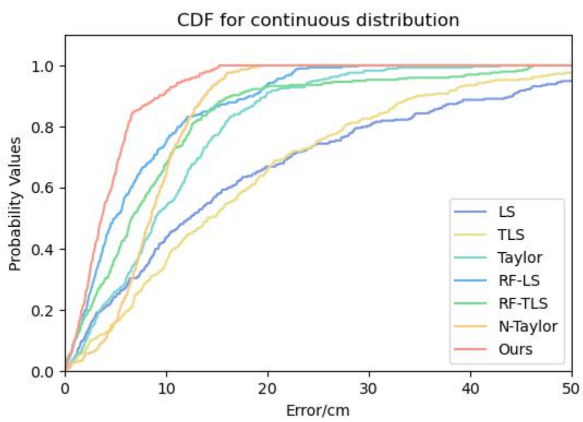

Fig. 3. CDF of localization errors sourced from different methods

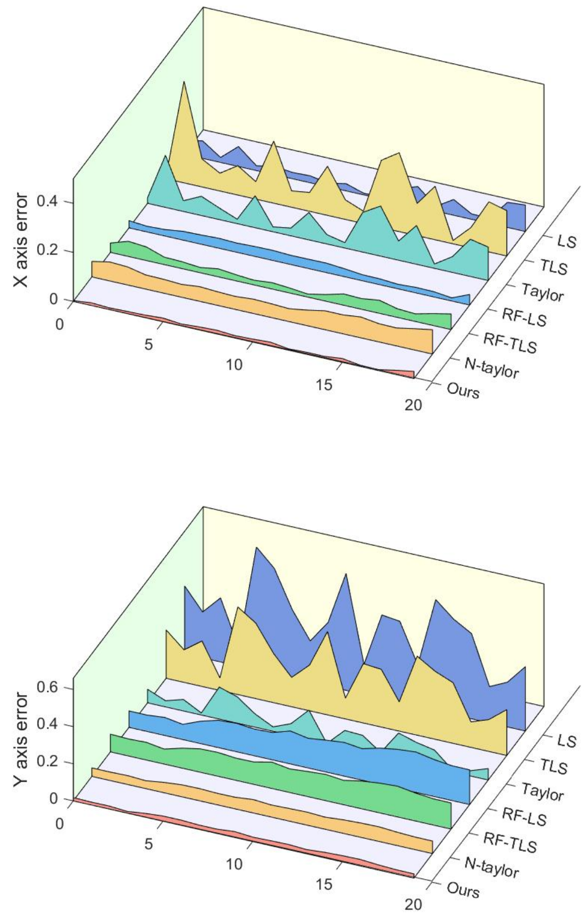

Fig. 4. Localization error distribution on two axes sourced from different methods

optimization methods, the localization error distribution of the two coordinate axes moving along the X-axis and Y-axis respectively is shown in Fig. 4. According to these results, the above localization methods have different localization precision on two axes. After using KF, the localization performance of Taylor, LS, and TLS on the two axes is significantly better than that of the corresponding method without filtering, indicating that KF plays an important role. When two methods using Taylor algorithm (UWB-RF and Taylor) are used, the average localization errors before and after filtering are 1.10 cm and 8.74 cm on the X-axis, respectively, and the localization accuracy is improved by 87.41%. The average localization errors on the Y-axis are 2.43 cm and 7.29 cm, respectively, and the localization accuracy is improved by 66.67%. The average error of RF-TLS method on X-axis and Y-axis is 3.59 cm and 12.73 cm respectively, and the accuracy of UWB-RF method is 69.36% and 80.91% higher than RF-TLS method. The average error of RF-LS method on X-axis and Y-axis is 2.92 cm and 15.60 cm respectively, and the accuracy of UWB-RF method is 62.33% and 84.42% higher than RF-LS method. It can be concluded that UWB-RF combines RF algorithm, KF and Taylor algorithm, making it become an effective robust optimization approach with significant advantages compared with other methods.

5. CONCLUTION

In order to eliminate the influence of NLOS error in complex indoor senarios, a new localization system which is named as UWB-RF system has been proposed. Taylor algorithm has been applied to improve the positioning accuracy for the presented symstem. The method can improve the estimation accuracy of UWB localization system in complex underground environment. RF algorithm is used to identify whether LOS or NLOS is between AN and TN. UWB ranging information and identified environmental information are applied to the measurement update of KF. After filtering, Taylor algorithm is used to further optimize localization and distance information, so as to achieve high-precision indoor localization. Experimental results show that the estimation accuracy and robustness of this system are better than other methods, and the calculated trajectory is closer to the actual trajectory than other methods, showing excellent localization performance.

6. ACKNOWLEDGMENT

This work is supported by the National Key Research and Development Program (2018YFB2100301); National Natural Science Foundation of China (61972131).

7. REFERENCES

[1] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE signal processing magazine*, vol. 22, no. 4, pp. 70–84, 2005.

[2] J. Zhang and C. Shen, "Research on uwb indoor positioning in combination with tdoa improved algorithm and kalman filtering," *Modern Electronics Technique*, vol. 39, no. 13, pp. 1–5, 2016.

[3] J. Schroeder, S. Galler, K. Kyamakya, and K. Jobmann, "Nlos detection algorithms for ultra-wideband localization," in *2007 4th Workshop on Positioning, Navigation and Communication*. IEEE, 2007, pp. 159–166.

[4] V. Savic, E. G. Larsson, J. Ferrer-Coll, and P. Stenumgaard, "Kernel methods for accurate uwb-based ranging with reduced complexity," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1783–1793, 2015.

[5] X. Shi, Y. H. Chew, C. Yuen, and Z. Yang, "A rss-ekf localization method using hmm-based los/nlos channel identification," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 160–165.

[6] C. Wang, A. Xu, J. Kuang, X. Sui, Y. Hao, and X. Niu, "A high-accuracy indoor localization system and applications based on tightly coupled uwb/ins/floor map integration," *IEEE Sensors Journal*, vol. 21, no. 16, pp. 18 166–18 177, 2021.

[7] M. Heidari and K. Pahlavan, "Identification of the absence of direct path in toa-based indoor localization systems," *International Journal of Wireless Information Networks*, vol. 15, no. 3, pp. 117–127, 2008.

[8] Y. Yasyukevich, A. Mylnikova, and A. Vesnin, "Gnss-based non-negative absolute ionosphere total electron content, its spatial gradients, time derivatives and differential code biases: bounded-variable least-squares and taylor series," *Sensors*, vol. 20, no. 19, p. 5702, 2020.

[9] A. Ribeiro, I. D. Schizas, S. I. Roumeliotis, and G. Giannakis, "Kalman filtering in wireless sensor networks," *IEEE Control Systems Magazine*, vol. 30, no. 2, pp. 66–86, 2010.

[10] N. Modalavalasa, G. S. B. Rao, K. S. Prasad, L. Ganesh, and M. Kumar, "A new method of target tracking by ekf using bearing and elevation measurements for underwater environment," *Robotics and Autonomous Systems*, vol. 74, pp. 221–228, 2015.

[11] J. Wang, T. Zhang, X. Xu, and Y. Li, "A variational bayesian based strong tracking interpolatory cubature kalman filter for maneuvering target tracking," *IEEE Access*, vol. 6, pp. 52 544–52 560, 2018.

[12] H. Wymeersch, S. Maranò, W. M. Gifford, and M. Z. Win, "A machine learning approach to ranging error mitigation for uwb localization," *IEEE transactions on communications*, vol. 60, no. 6, pp. 1719–1728, 2012.

[13] K. K. Cwalina, P. Rajchowski, O. Blaszkiewicz, A. Olejniczak, and J. Sadowski, "Deep learning-based los and nlos identification in wireless body area networks," *Sensors*, vol. 19, no. 19, p. 4229, 2019.

[14] S. Krishnan, R. X. M. Santos, E. R. Yap, and M. T. Zin, "Improving uwb based indoor positioning in industrial environments through machine learning," in *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE, 2018, pp. 1484–1488.

[15] W. Zhang, C. Wu, H. Zhong, Y. Li, and L. Wang, "Prediction of undrained shear strength using extreme gradient boosting and random forest based on bayesian optimization," *Geoscience Frontiers*, vol. 12, no. 1, pp. 469–477, 2021.

[16] D. Feng, C. Wang, C. He, Y. Zhuang, and X.-G. Xia, "Kalman-filter-based integration of imu and uwb for high-accuracy indoor positioning and navigation," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3133–3146, 2020.

[17] Y. Cheng and T. Zhou, "Uwb indoor positioning algorithm based on tdoa technology," in *2019 10th international conference on information technology in medicine and education (ITME)*. IEEE, 2019, pp. 777–782.

[18] C.-D. Wann and C.-S. Hsueh, "Nlos mitigation with biased kalman filters for range estimation in uwb systems," in *TENCON 2007-2007 IEEE Region 10 Conference*. IEEE, 2007, pp. 1–4.

[19] M. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 10, pp. 1440–1454, 1987.