Deep Learning based Wireless Human Motion Tracking for Mobile Ship Environments

Kezhong Liu, Wen Yang, Mozi Chen, Kai Zheng, Xuming Zeng, Shengkai Zhang and Cong Liu

***Abstract*—Being able to track passengers' movement without invasion of their privacy plays an important role in cruise ships; it enables crucial location-based services such as maritime search&rescue, tourist services, and epidemic prevention. The past few years have witnessed commodity WiFi holding great potential that provides such services available thanks to its ubiquitous in indoor scenarios. However, existing WiFi-based tracking methods suffer from huge performance degradation in sailing ships due to their complex metal structures and dynamic hull deformation caused by engines and waves/payloads pressure. In this paper, we present CRLoc, a deep learning-based passive human tracking system that can overcome the practical limitations of traditional WiFi-based localization approaches applied in a multipath-rich and mobile ship environment, and provide decimeter-level tracking accuracy in cruise ships. Specifically, we make two contributions, i.e, we propose a super-resolution parameter estimation algorithm that better characterizes ship indoor environments, and a deep neural network-based end-to-end solution to remove the impact of noise, interference, and mobility in ships. The real-world implementation and extensive experiments in several passenger ships demonstrate that CRLoc tracks human motions with a median error of 92 cm, better than state-of-the-art localization methods. To our knowledge, this is one of the first WiFi-based passive human motion tracking system in a cruise ship environment.**

***Index Terms*—WiFi Localization, Mobile Ship Environment, Super-resolution Algorithm, Convolutional Recurrent Neural Network**

I. INTRODUCTION

Being an essential part of the emerging intelligent ships, crew and passengers' location perception and trajectory recording plays an important role in ship management and safety insurance. A recent tragedy, such as the outbreak of COVID-19 on Diamond Princess in 2020 [1], illustrates the serious consequences of the lack of accurate movement information of passengers in such an enclosed space would lead to chaotic management. A ubiquitous and device-free human motion tracking localization system can be used to prevent such incidents, which can also be utilized in maritime search&rescue, tourists services, and epidemic prevention [2].

K. Liu is with the School of Navigation, Wuhan University of Technology, Wuhan 430063, China, National Engineering Research Center for Water Transport Safety, Wuhan 430063, China and Hubei Key Laboratory of Inland Shipping Technology, Wuhan 430063, China (e-mail: kzliu@whut.edu.cn).

W. Yang, M. Chen, K. Zheng, X. Zeng, are with the School of Navigation, Wuhan University of Techology, Wuhan 430063, China (e-mail: {wenyang, chenmz, kai, zengxvming}@whut.edu.cn).

S. Zhang is with the School of Information Engineering, Wuhan University of Techology, Wuhan 430063, China (e-mail: shengkai@whut.edu.cn).

C. Liu is with the Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521, USA (e-mail: congl@ucr.edu)

* The first three authors are co-first authors and have equal contributions.

* Corresponding author: Mozi Chen (chenmz@whut.edu.cn).

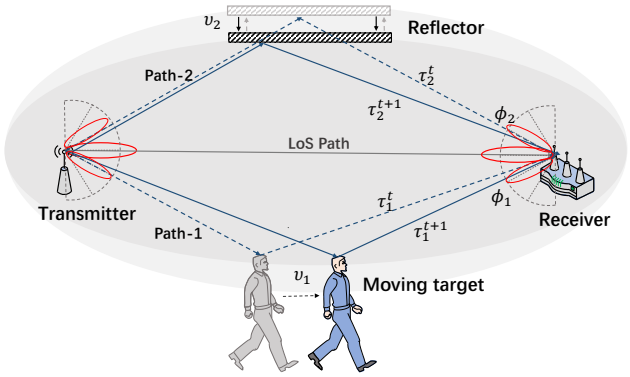


Fig. 1. The principle of WiFi-based device-free localization and tracking through. We obtained the time-of-flight (ToF) τ, angle-of-arriver (AoA) ϕ and doppler frequency shift (DfS) υ from WiFi signal. The reflectors such as deformed hulls in a cruise ship changed the propagation of wireless signals and interfered with channel parameters.

In the last few years, we have seen a lot of emerging techniques for passive human indoor localization, such as computer vision [3] [4], ultra-wideband (UWB) [5] [6], radio-frequency identification (RFID) [7] [8] and WiFi, etc. The development of the first three techniques raises a number of concerns, including privacy concerns, deployment costs, and intrusive equipment. In contrast, the WiFi-based approach has become more popular owing to its ubiquitous availability and non-line-of-sight characteristics [9]. Furthermore, the emerging WiFi tools such as the latest 802.11ax protocol channel state information (CSI) acquisition tool can provide 256 subcarriers and 160MHz bandwidth, which strongly supports WiFi-based smart applications [10]. Basically, WiFi-based localization systems are generally divided into two folds: *fingerprint-based* and *model-based* schemes. A fingerprint-based scheme typically analysis the statistical characteristic of the wireless signal propagation at each location and performs similarity matching between real-time and pre-collected data, which obviously is labour-intensive and the accuracy is vulnerable to environmental changes. To overcome this shortcoming, a model-based scheme that does not require a site-survey for fingerprints collection, thereby avoiding time-consuming dataset collection. As shown in Fig. 1, current advanced systems typically extract location-related wireless channel parameters from the WiFi signal, such as the time-of-flight (ToF), angle-of-arrival (AoA) and Doppler frequency shift (DfS), and attenuation of the incident signal. Using these parameters, we can determine the location of people in a space with considerable accuracy.

Unfortunately, existing wireless localization or tracking systems suffer from huge accuracy degradation in a mobile ship environment because they cannot address the impact of dynamic characteristics and metal reflection. To understand the impact factors of a mobile ship, our early attempts are to statistically quantify the relationship between random ship motion and CSI changes based on data-driven methods [11]. However, studies on the relationship between time-varying ship environments and wireless signals from a model-based perspective are scarce. Based on our observations, in a mobile ship environment, the wireless signal arrives at the receiver along different paths at each moment, and the number of paths changes dynamically, which is different from the general environment. This phenomenon may be a result of the inevitable elastic hull deformations caused by the internal and external stress from the load, waves, and engines [12]. This characteristic reduces the accuracy of channel parameter estimation because the number of multipaths needs to be specified. In addition, owing to the steel material of the ship, high-order reflected signals and other wireless signal sources cause severe interference to the target signal and decrease parameter precision. Moreover, the human body can be confused with other reflective objects owing to a strong metal shot. To achieve high-precision localization and tracking in a ship environment, these challenges must be overcome.

In this paper, we propose a deep learning-based human motion tracking system that can accurately track the location of people in a mobile ship environment. The system was designed to 1) consider the dynamic change of multipath number in the ship environment and realize high-precision location-related parameter estimation from WiFi signal, 2) overcome the impact of high noise in the ship and achieve acceptable localization accuracy, and 3) realize the continuous localization and tracking of people using existing commercial WiFi facilities. The proposed system, CRLoc, is a WiFi-based system that can accurately localize and track a moving human and overcome the limitations of the mobile ship environment. The design consists of two main parts. First, based on the CSI Motion model [13], an improved super-resolution parameter estimation algorithm, namely the sparse variational bayesian space-alternating generalized expectation-maximization algorithm (VBSAGE), is proposed to extract high dimension channel parameters from the CSI signal. VBSAGE was used because it can simultaneously estimate the number of components in a mixture and their parameters, which adapts to the dynamic environment of the ship. Second, we reorganized the estimated parameters to synthesize the distance-direction likelihood profile of the target and designed a modified convolutional recurrent neural network (CRNN) model to obtain the target location. The network can automatically learn the refined parameters of the target from severe noise and distinguish the moving human from numerous reflections. Real-world implementation and extensive experiments demonstrate that CRLoc tracks human motions with a median error of 92 cm.

Our main contribution are as follow:

- We propose a novel super-resolution parameter estimation approach to simultaneously estimate the number of multipaths and their parameters, which reduces the accuracy degradation caused by the dynamic multipath in a mobile ship environment;
- We present a deep learning end-to-end solution to overcome the agnostic caused by severe noise and dynamic multipath scenario. Specifically, a two-dimensional (2D) likelihood profile was generated using multipath parameters, and a CRNN model was proposed to automatically learn the refined target parameters and extract the person's trajectory.
- CRLoc was implemented and evaluated with commodity Intel 5300 Wi-Fi cards. Extensive experiments on a real world ship demonstrated high accuracy and significant performance improvement for a mobile ship environment.

II. RELATED WORKS

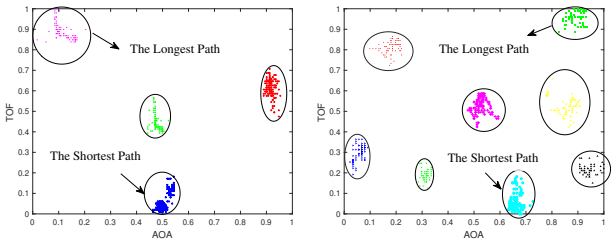
Wireless-based passive localization and tracking have drawn considerable attention in recent years. In this section, we briefly review state-of-the-art localization work based on different wireless technologies. Subsequently, we introduce the most relevant WiFi-based indoor localization systems.

A. State-of-the-art Indoor Localization.

In recent years, a large number of indoor localization solutions have emerged. Commonly used indoor localization technologies include RFID, UWB, Computer vision, Bluetooth, acoustic, Zigbee, and WiFi. BERT-ADLOC [14] integrate Bluetooth with state-of-the-art technologies, such as cloud computing and deep neural networks, to achieve decimeter-level localization. 3D-LBMS [15] combines BLE and multiple sensors to achieve meter level 2D localization accuracy and submeter level 3D altitude estimation accuracy. RFind [7] leverages the underlying physical properties of RFID to emulate a large bandwidth and uses it for localization. MRL [8] studies the problem of tag localization using RFID-augmented robots, which are used for automatic item fetching and misplacement detection in warehouses. Smart metasurfaces [5] are considered an important auxiliary technology for UWB high-precision localization in complex indoor environments and have received widespread attention recently. iVR [3] integrates vision and radio localization systems and achieves sub-meter accuracy with indoor semantic maps, which are automatically generated from only two surveillance cameras. A recent study [4] further enhanced indoor localization with multimodal sensing using camera images, inertial measurement unit sensors and WiFi signals on a smartphone. However, as mentioned earlier, we did not adopt these technologies for several reasons.

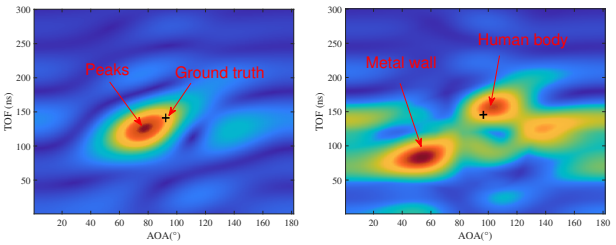
B. WiFi-based Localization and Motion Tracking.

WiFi-based localization systems typically use CSI and received signal strength (RSS) signals. In recent years, CSI-based passive human localization and tracking have received great attention because CSI can provide higher fine-grained channel information than RSS. These strategies can be divided into two categories: fingerprint-based methods and physical



(a) Few Multipath. (b) Rich Multipath.

Fig. 2. Clusters of AoA and ToF in two adjacent time periods, which describes the number of multipaths. This tells us that the number of multipaths is dynamically changing. (a) An example of a small number of multipaths, only 4 multipaths. (b) An example of a large number of multipaths, including 9 multipaths.



(a) Wide peaks. (b) Multipath peaks.

Fig. 3. Effect of signal in a ship environment. (a)The wide peak due to severe environmental noise. (b)The multipath peaks and the groundtruth (red cross) isn't at the largest peak.

model-based methods. AutoFi [16] proposes an approach to automatically calibrate the localization profiles in an unsupervised manner and solve the inconsistency between fingerprints and new profiles when the environment changes. Fidoar [17] uses a variational autoencoder to localize new users and trains a domain-adaptive classifier to remove the disturbance of environmental changes on CSI. Chronos [18] proposes a channel splicing algorithm that combines 2.4G and 5G Wi-Fi channels and improves ToF resolution to sub-nanometer level. MaTrack [19] consists of a novel dynamic-MUSIC method to extract the object reflection path and estimate the AoA for device-free localization. Spotfi [20] uses the 2D-MUSIC algorithm to jointly estimate ToF and AoA, which improves the accuracy of the parameters because the parameters are more separable in two dimensions than in one dimension. In some studies [21] [13] [22], different parameter combinations of ToF, AoA, angle-of-departure (AoD), DfS and signal attenuation are extracted to achieve accurate channel parameter estimation.

III. PROBLEM FORMULATION

The goal of this study is the development of a reliable WiFi-based indoor localization system in a special ship environment. However, compared with an ordinary indoor environment, the ship environment has a severe impact on the wireless signal. In this section, we first introduce the basic CSI model and describe the relationship between CSI measurements and human's movement. Next, we present three practical challenges which we need to overcome in real-world sailing ships.

A. Signal Superposition Model

We determine the target location based on the wireless signal reflected by the target, which contains key information, such as distance, direction, frequency shift, and signal strength. However, a number of reflectors in the room can reflect wireless signals. In reality, the signals from all reflectors superpose together at sensors. To capture this fact, we modeled each CSI measurement as the sum of the received signals from all reflectors, which is referred to as a signal superposition model.

For simplicity, we provide a simplified version of the model mentioned here, and the CSI measurement Z can be expressed as follows:

$$Z = \sum_{l=1}^{L} \alpha_l s(\theta_l) + \xi \tag{1}$$

where α_l denotes the complex attenuation between the transceivers and $s(\boldsymbol{\theta}_l)$ denotes the phase shift function of the parameter θ of the lth path. Here $\theta = [\tau, \phi, \upsilon]$, and τ, ϕ, υ represent ToF, AoA, DfS, respectively. ξ represents the additive background noise. L is the total number of multipaths, which is typically constant in a normal indoor environment. Here we briefly introduce the target parameters in the model.

- *Time of Flight (ToF).* ToF is the signal propagation time from the transmitter to the receiver. By multiplying the ToF by the speed of light, we can get the distance from the target to the transceiver.
- *Angle of Arrival (AoA).* AoA means the direction of incoming signal at the receiver. It reflects the orientation of the target relative to the antenna array.
- *Doppler frequency Shift (DfS).* DfS represents the path length change caused by the target movement at different times, which can reflect the target speed information.

The above parameters represent the distance, direction and speed information of the target, and can be calculated by the parameter estimation method. However, we encounter some challenges in a dynamic environment of the ship.

B. Practical Challenge

Our study aims to address the significant shortcomings of the current localization systems applied to mobile cruise ships. Here we present three practical challenges that affect localization accuracy as follows:

1) Dynamic multipath number: The multipath effect becomes a more troublesome issue in a mobile steel ship environment because of its dynamic nature. This has two aspects. On the one hand, the ship environment is time-varying because of ship deformation [11]. Similar to a moving human, deformed steel walls, floors, and ceilings will also cause a phase change and frequency shift of the reflected signal. This is different from the normal environment, hindering the complete removal of harmful multipaths using a filter [13] or the difference

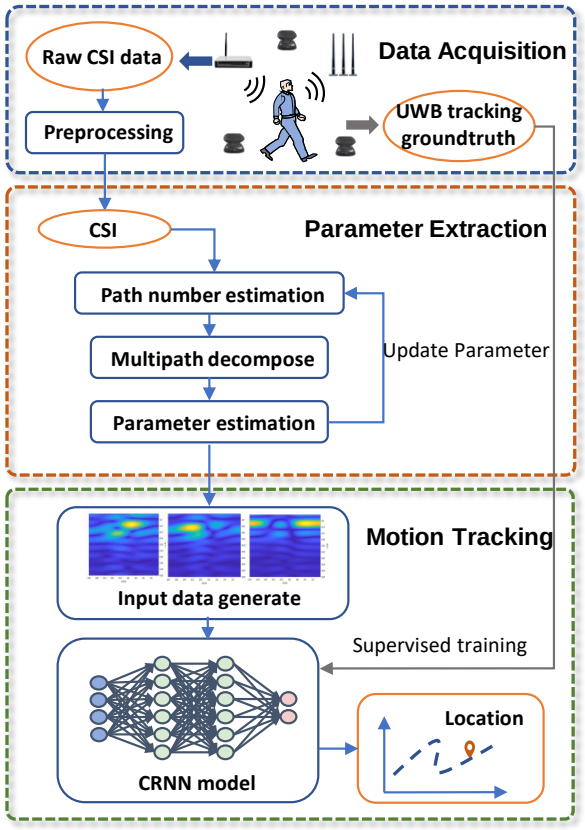


Fig. 4. System Overview.

between the two measurements [23]. Therefore, we estimated all multipaths. On the other hand, owing to the varying degrees of deformation, the number of multipath is also dynamic, as shown in Fig. 2. In contrast, it is typically considered that there are approximately five main multipaths in a normal indoor environment [20]. When we decompose the multipath and noise signals without knowing the number of multipaths, the actual signal on each multipath cannot be accurately restored, which will eventually lead to inaccurate channel parameter estimation.

2) Strong metal reflection: The wireless signal experiences high noise interference in such a steel enclosed space, including high-order reflected residual signals and abundant wireless equipment. Because steel has a stronger reflection on wireless signals, it undergoes higher-order reflections in the cabin of the ship. This poses a challenge for the implementation of accurate indoor localization technologies. Specifically, when we perform location feature extraction, more noise signals are incorporated into the CSI on each multipath, leading to an inaccurate estimation. Fig. 3 (a) shows the deviation of the estimated peak from the true location in the ToF-AoA likelihood profile. Using the CSI collected in the ship environment, the heat map of the estimated feature will have a wider peak, making it more difficult to obtain the correct feature.

3) Distinguish reflected objects: In our design, to make parameter estimation more accurate, we plan to estimate the channel parameters of the signal from all reflectors and not only the target of interest. However, as shown in Fig. 3 (b),the parameters of irrelevant reflectors cause confusion in target locations, hindering the continuous tracking of people. In previous studies, the strongest reflected signal was believed to be the signal reflected by a human. After removing the static multipath from the static object, only the signal reflected by moving people remains. In our study, the reflected signal from irrelevant objects is dynamic and cannot be neglected, particularly when people are close to them. Therefore, we must design a plan that can effectively distinguish and track moving humans.

IV. OVERVIEW

The aim of our study was to enable WiFi-based device-free indoor human tracking, considering the particularity of the ship environment. As shown in Fig. 4, CRLoc comprises two key components: *location-related parameter extraction module* and *CRNN-based motion tracking module*.

1) Parameter Extraction: The parameter extraction module primarily aims to estimate the location-related parameters of the dynamic multipath signals in a mobile ship environment. To address this issue, we used the proposed VBSAGE algorithm, which jointly estimates the number of multipaths and their corresponding parameters [24]. Upon receiving the preprocessed CSI series, the VBSAGE algorithm first sets the multipath number, L, a rough value and decomposes the clustered CSI into L components. The maximum likelihood is used to estimate the location-related parameters. Subsequently, the estimated parameters were used to refine the multipath number and iterate until it converges. Finally, we obtain the location-related parameters of all $\hat{L}$ multipath.

2) Motion Tracking: The motion-tracking module implements a modified CRNN for human motion tracking. With the multipath parameter series as input, CRLoc first transposes it into a 2D ToF-AoA likelihood profile. In this manner, the problem is transformed into an issue similar to image processing and can be solved using powerful deep learning methods. Our model includes convolutional, recurrent, and fully connected (FC) modules. A convolutional module can extract spatial features and compress data [25]. The recurrent module can find the spatial features of human-related parameters from mixed parameters, and has proven to be suitable for location tracking [26]. Finally, the FC module maps these features onto the location coordinates of the person. In principle, CRLoc achieves accurate human localization and tracking, which requires only one-time training of the CRNN network and can be directly adapted to many new ships.

V. PARAMETER ESTIMATION USING VBSAGE

In this section, we introduce a method to extract location-related parameters from CSI signals using the proposed sparse VBSAGE algorithm. First, we consider a simplified scenario in which signal travels along only one path to the receiver.

In sequence, we propose a sparse joint estimation method to address the dynamic multipaths and parameter estimation.

A. Basic Ideal

Our purpose is to estimate the parameters related to the target location from the CSI measurement Z, considering the dynamics of the ship environment. However, in a mobile ship environment, the dynamic wireless channel has significant impact on the accuracy of parameter estimation. Intuitively, when we decompose the CSI of various paths that are mixed together, if the number of components is greater than the actual number of paths, the CSI of each path will be smaller than the real situation, and vice versa. Our concept consisted of developing an adaptive algorithm to estimate the paths and their corresponding channel parameters and achieve a joint estimation of the number of paths and channel parameters.

For simplicity, consider a specific propagation path l, where the CSI measurement x_l on this path is follows:

$$x_l = \alpha_l s(\theta_l) + \xi_l \tag{2}$$

where $s(\theta_l) = e^{-j2\pi\Delta f_j \tau_l} e^{-j2\pi f_c \Delta d_k \phi_l / c} e^{j2\pi \upsilon_l \Delta t_i}$. $s(\theta)$ consists of three parts: the phase shift caused by ToF τ, AoA ϕ, DfS υ on frequency, space, and time diversity, respectively. $\Delta f_j, \Delta d_k, \Delta t_i$ indicate the frequency, distance and time differences between the current subcarrier and the first subcarrier, the current antenna and the first antennas, the current packet and the first packet, respectively. f_c represents the carrier frequency of the channel and c represents the speed of light. α_l represents the complex attenuation from the transmitting antenna to the receiving antenna along the lth path. A additive Gaussian white noise component ξ_l is obtained by arbitrarily decomposing the total noise ξ, $\Sigma_l = \mathbb{E}\left\{\xi_l \xi_l^H\right\} = \beta_l \Sigma$. β_l is the decomposition factor of the lth component and $0 \leqslant \beta_l \leqslant 1$.

As the VBSAGE algorithm still belongs to the generalized EM algorithm family, the process of channel parameter estimation is also divided into expectation and maximization steps. Let $\hat{x}$ and $\hat{x}'$ denote the estimation of the CSI measurement x in the current iteration and last iterations. Suppose there are L paths in one iteration, that is, the CSI observation is divided into L components. In the expectation step, we restore the unobservable CSI measurements on each path. When observation data Z and the last estimated parameter $\hat{\theta}'$ are provided, we obtain $\hat{\boldsymbol{x}}_l$ by calculating the conditional expectation of x as follows:

$$\hat{\boldsymbol{x}}_l = \boldsymbol{x}_l'(m; \hat{\theta}_l') + \beta_l [Z(m) - \sum_{l=1}^{L} \boldsymbol{x}_l'(m; \hat{\theta}_l')] \tag{3}$$

where $m = (i, j, k)$ denotes the data at time i, subcarrier j and antenna k.

After obtaining the CSI on each path, VBSAGE obtains the channel parameters by using maximum likelihood estimation for each component, expressed as follows:

$$\begin{aligned} P_l &= \log p(\hat{\boldsymbol{x}}_l | \theta_l, \hat{\alpha_l}) - \hat{\Phi} s(\theta_l)^H {\Sigma_l}^{-1} s(\theta_l), \\ \hat{\boldsymbol{\theta}}_l &= \arg\max_{\theta_l} P_l \end{aligned} \tag{4}$$

In (4), P_l is the ToF-AoA likelihood profile of the lth path. The previous term is the same as the basic SAGE algorithm, and the latter term is a regularization term, with the posterior variance Σ_l of α_l acting as a regularization constant. Note that $\theta_l = (\tau_l, \phi_l, \upsilon_l)$, and we use gradient descent method to update $\tau_l, \phi_l and \upsilon_l$. Specifically, we fix these parameters except one and search for the value of the unfixed parameter.

B. Sparse Multipaths Number

In this section, we set out to solve the dynamic changes in the path number of parameter estimation. This problem is called the model order selection problem and has been discussed in many studies [27] [28]. To implement model order selection during a parameter update iteration, we introduce a sparse coefficient w related to attenuation α, which obeys the gamma distribution. The VBSAGE algorithm tests three types of sparsity priors. In this study, we selected the Gaussian sparsity prior $p(\alpha_l | w_l) = \mathbb{CN}(\alpha_l; 0, w_l^{-1})$, and then updated α_l as follows:

$$\begin{aligned} \hat{\Phi}_l' &= (\hat{w}_l + s(\theta_l)^H {\Sigma_l}^{-1} s(\theta_l))^{-1}, \\ \hat{\alpha}_l' &= \hat{\Phi}_l' s(\theta_l)^H {\Sigma_l}^{-1} \hat{x}_l \end{aligned} \tag{5}$$

This formula is a regularized least-squares estimate of ${\hat{\alpha}_l}'$ given $\hat{x}_l$ and $\hat{\theta}_l$. Because the sparse parameter w obeys the gamma distribution, we can update it as follows:

$$\hat{w}_l' = \frac{1}{|\hat{\alpha}_l'|^2 + \hat{\Phi}_l} \tag{6}$$

To control the signal component, the VBSAGE algorithm sets a pruning condition. Assuming that the sparse parameter w is always greater than zero, a simplified pruning condition can be obtained as follows:

$$\left| s(\hat{\theta}_l)^H {\Sigma_l}^{-1} \hat{x}_l \right|^2 > s(\hat{\theta}_l)^H {\Sigma_l}^{-1} s(\hat{\theta}_l) \tag{7}$$

When (7) is not satisfied, the sparse parameter w_l approaches infinity, and thus the attenuation in (4) equals zero. This allows us to sparse multipath number during a parameter update iteration, that is, joint multipath number detection and parameter estimation.

C. Joint Sparse Multipaths and Parameter Estimation

Algorithm 1 summarizes the main steps of the parameter estimation using the sparse VBSAGE algorithm. For moment, we assume that at one iteration, the parameters, x_l, θ_l, α_l and ω_l, $l \in \{1, 2, ...\hat{L}\}$ are known for $\hat{L}$ paths. In the next iteration, we estimate each parameter individually one in a certain order, and the values of the other parameters are fixed while estimating one parameter. The update iteration was repeated for all paths until the number of paths and their parameters converge. The number of paths might be reduced during one update iteration: at each round, the updated multipath component undergoes a test specified by 7. When the condition is not satisfied, the corresponding path is not considered as a real physical wireless path and is removed. Note that during initialization, L is set to an appropriately large value L_{max}.

Algorithm 1 Parameter estimation using Sparse VBSAGE

Input: : CSI Measurement $Z(m)$
Output: Location-related parameter $\Theta = (\theta_l)_{l=1}^{\hat{L}}$

```
Initialization. L = L_max, Θ = Θ_0
while ||Θ'' − Θ'|| > ξ && L̂' ≠ L̂ do
    for l = 1 → L do
        Update x_l from (3)
        Update θ_l from (4)
        if Condition (7) is true then
            Update α_l from (5)
            Update ω_l from (6)
            L ← L̂
        else
            Remove the lth path; L ← L̂ − 1
        end if
    end for
end while
```

Initialization. The original VBSAGE provides a simple bottom-up initialization strategy that allows us to infer the initial variational parameters from the measurement $Z(m)$ by starting with an empty model. Considering the running time of the model, we used a simple initialization method, assuming all variational parameters to be zero.

VI. CRNN-BASED LOCALIZATION

Although the channel parameters of all multipaths are obtained simultaneously using the proposed VBSAGE algorithm, we cannot completely trust them because of high noise in the ship environment. Furthermore, we must solve the confusion of the reflector location, that is, identifying the target parameters from the cluttering parameters and mapping them to the target location.

A. Input Data Structure

In Sec. V, we obtained the multi-dimensional location-related parameters. Among these parameters, ToF and AoA are critical because they represent distance and direction, respectively, whereas DfS is used to refine the auxiliary results. Because we use 100 packets within 0.1s in location estimation, DfS may introduce location aliasing and cause the estimated location to deviate [26]. In our study, we use the estimated DfS to correct input data. Specifically, in a specified estimation, when the algorithm iterates N times and converges, we let the estimation algorithm iterate one more round, that is, execute formulas (3) and (4) again with a fixed value of DfS in the last iteration, expressed as follows:

$$P = \sum_{l=1}^{\hat{L}} \mathrm{VBSAGE}_{N+1}(\tau_l, \phi_l, \hat{\omega}_l, \hat{\alpha}_l, \hat{L}) \tag{8}$$

Where P represents the ToF-AoA likelihood profile for all path. The value of ToF τ_l and AoA ϕ_l range from 0 to 360 ns and from 0 to 180°, respectively. $\hat{\omega}_l$, $\hat{\alpha}_l$ and $\hat{L}$ represent the value of ω_l, α_l and L in the Nth iteration. In this manner, we turn the problem of finding the correct parameters into a problem of image processing, and we can then use methods in the field of image processing to solve this problem. In sequence, we select the promising CRNN model [29] to address the problem mentioned in Sec. III. Our key insight is that the location of human changes continuously and will not change too much over a short time. Furthermore, the motion pattern of human is different from the other reflectors such as deformed bulkheads. Next, we introduce the design details of our CRNN model.

B. Network Architecture

The CRNN model has achieved considerable success in the task of optical character recognition. Recent studies have also applied this model to wireless signal direction estimation [30]. Our model differs from the original CRNN model, which is primarily composed of CNN, long short-term memory (LSTM) networks, and FC modules. But they are similar in structure to a certain extent, so we still use this term. Fig. 5 shows the architecture of the proposed network.

Convolutional module. The convolutional module was used to extract the spatial features and compress the data. Owing to the high sparseness and preservation of the spatial locality of the ToF-AoA profile, CNN is capable of handling this task. Specifically, the input data P is a tensor of $M \times N \times T$ dimension, where M is the length, N is the width, and T is the number of profiles. The convolutional module contains three layers and each layer applies a set of 3×3 convolutional filters. The pooling layer further compresses the data by operating on the data in the window, for example, by selecting the largest element in the window as the output value in max-pooling. Here we conducted max-pooling over 2×1 windows for the first layer and 2×2 windows for the remainder. After three convolution and pooling operations, we flattened the output matrix into a vector V, such that it could be used as the input of the following recurrent layers for temporal modeling. Given input data P, we can obtain the feature representation as follows:

$$V = \mathrm{CNN}(P) \tag{9}$$

Recurrent module. For the RNN module, we used LSTM to identify the parameters of interest, which can effectively capture the context characteristics of a time series. A common LSTM unit is composed of cell c, input gate i, output gate o and forget gate f. The cell remembers the values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell. The cell state c contains the location context information and feeds into the current loop network after concatenating with input data. It expressed as follows:

$$\begin{pmatrix} i \\ f \\ o \\ j \end{pmatrix} = \begin{pmatrix} \sigma_L \\ \sigma_L \\ \sigma_L \\ \tanh \end{pmatrix} W_L \binom{h_{t-1}}{V_t}$$
$$c_t = f \odot c_{t-1} + i \odot g \tag{10}$$
$$h_t = o \odot \tanh(c_t)$$

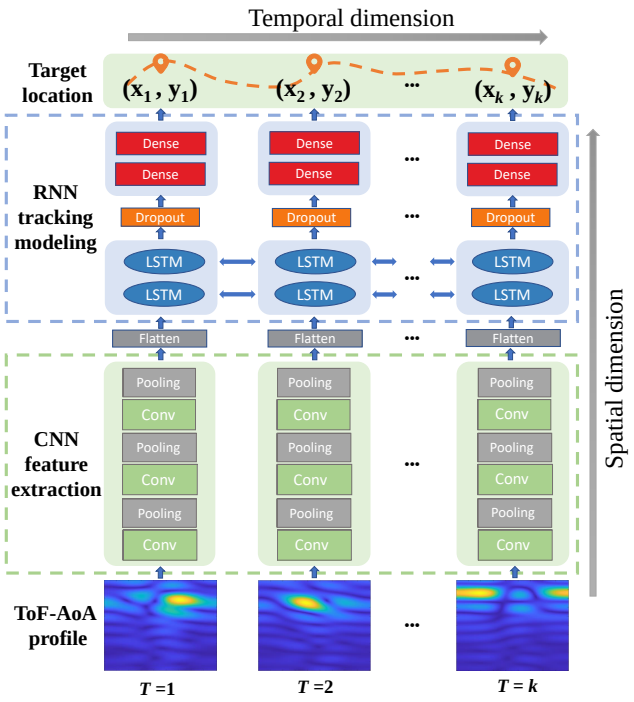


Fig. 5. Architecture of the network.

Here, $\odot, \sigma_L$ and $\tanh$ represent the element-wise multiplication, logistic sigmoid activation, and hyperbolic tangent function respectively.

We used two stacked bidirectional LSTMs recurrent networks to extract the interesting sequence information. In the training process, a discard rate of 0.1 was used to prevent overfitting. Next, the features of the LSTM layer were fed to the full connection layer. All operations in this part can be simply expressed as follows:

$$H = \mathrm{LSTM}(V) \tag{11}$$

Fully connected layers. The output features from the LSTM are fed into a FC layer. The input layer consists of 512 units. The output layer contains two units, which represent the X-axis and Y-axis coordinates. The core operation of the FC layer is the matrix-vector product, which transforms data from one feature space to another feature space linearly. This step is expressed as follows:

$$L = \mathrm{FC}(H) \tag{12}$$

Here $L = (x, y)$ denotes the target location, where x and y represent coordinates. The parameter settings of the network are summarized in Table 1.

C. Training and Localization

In this section, we explain the offline training and online localization processes. CRLoc extracts spatial features and compressed data through the CNN module, extracts temporal features through the LSTM module, and maps them to location coordinates using the FC layer. We simply express this process and the loss function as follows:

Table 1 Network configuration summary. 'c' ,'k', 's' and 'p' stand for channel number, kernel size, stride and padding size respectively.

Network type	Parameter set
Input	361×181 likelihood profile
Convolution	c:16, k:3×3,s:2×1,p:1
Maxpooling	Window:4×4,s:4
Convolution	c:32, k:3×3,s:1,p:0
Maxpooling	Window:4×4,s:4
Convolution	c:64, k:3×3,s:1,p:0
Maxpooling	Window:4×4,s:4
Flatten	-
Bidirectional-LSTM	#hidden units:512
Bidirectional-LSTM	#hidden units:512
Dense	#hidden units:512
Dense	#hidden units:2

$$\begin{aligned} V &= \mathrm{CNN}(P) \\ H &= \mathrm{LSTM}(V) \\ L &= \mathrm{FC}(H) \\ Loss &= \tfrac{1}{K}\textstyle\sum_{i=1}^{K}(l_i - g_i)^2 \end{aligned} \tag{13}$$

Here, l_i is the ith location estimated by CRLoc and g_i is the corresponding ground truth location. K is the step size of the LSTM module, which was set as 20. We used a simple minimum mean square error loss function as a training objective, because a simple Euclidean distance was sufficient to indicate the location difference between two points in the ordinary plane coordinate system. Location estimation is defined as the task of estimating whether each location on a predefined grid corresponds to the ToA–AoA likelihood profile. The tracking area was 6 m and we defined the resolution as 0.1 m, which resulted in 3,600 different combinations. During model training, we set the epoch to be 600 and the batch size to be 20. The base learning is set to 0.1 with a scheduler that reduces the learning rate by factor of 0.8 every 100 epochs. We employed the Adam optimizer with default parameters.

In the online localization stage, the collected CSI is preprocessed and then input to the sparse VBSAGE module for location-related parameter estimation. After being converted into the ToF-AoA profile, it is input to the CRNN module, and the location coordinates of each point in the trajectory are output. The moving average method was used to smooth the result and connect each segment of the trajectory. The entire algorithm is summarized in Algorithm 2.

Algorithm 2 CRLoc System

```
Place the devices using our optimal configuration;
Perform initialization;
while TRUE do
    Fetch new CSI samples;
    Perform precondition steps;
    Derive the location-related parameters with the sparse VBSAGE algorithm;
    Generate ToF-AoA profiles as input data;
    Use CRNN to map the profile to the target location;
end while
```

VII. IMPLEMENTATION AND EVALUATION

In this section, we first introduce the data collection methodology and CSI preprocessing method. In sequence, we evaluate the performance of CRLoc against four state-of-art localization techniques. System effectiveness and certain influencing parameters are also discussed.

A. Experiments Setup

We implemented CRLoc using a transmitter with one antenna and a receiver with three antennas, where the antennas formed a linear array and the distance was half of the wavelength. The transmitter and receiver were a pair of off-the-shelf Thinkpad T400 laptops with an Intel 5300 NIC. The system was Ubuntu and Linux 802.11n CSI Tool [31] was installed to collect the CSI. The WiFi NIC was set to operate in monitor mode on channel 165 at 5.825 GHz, and the transmission rate of packets was set to 1000 Hz. We implemented CRLoc in MATLAB and Pytorch. The processing computer was equipped with an Intel i7-9750 CPU and an NVIDIA GTX1660Ti GPU.

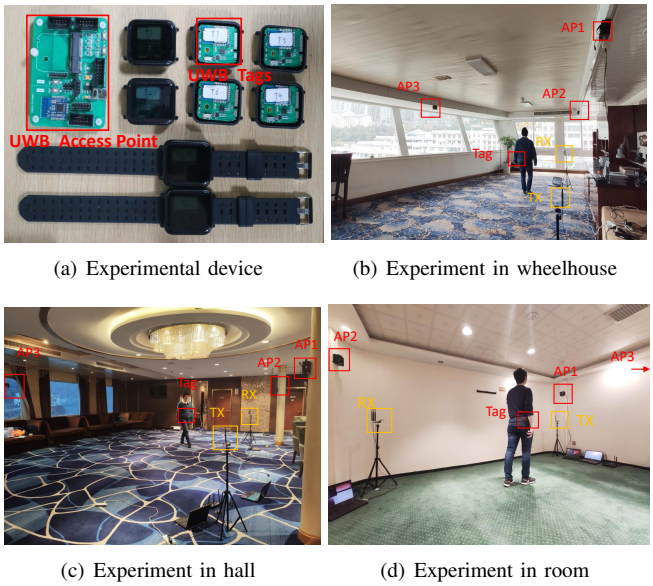
To fully verify the performance of CRLoc, we conducted extensive experiments in three indoor environments on the Golden Six cruise ship: an empty room, conference hall, and wheelhouse, which can represent the typical ship indoor environment. Fig. 6 shows the general environmental features and deployment of the devices under different scenarios. In total, four volunteers of different heights and weights participated in the experiment and walked along arbitrarily different shapes of trajectories such as lines, rectangles, and circles. Fig. 7 shows some examples of tracking results of CRLoc. We also tested the performance of the system, which are discussed later.

Data set. During the three-day voyage, we collected CSI data and the corresponding ground truth data for different periods. We collected a total of 8×10^6 packets, which were divided into 8×10^4 CSI segments, and each segment corresponds to a location to be estimated. Subsequently, every 20 positions formed a training trajectory, that is, a total of 4000 trajectories. As for the ground truth, we used advanced UWB-based tracking solutions, which had a localization error of only 10 cm owing to their wide bandwidth. Specifically, three UWB APs were installed in the tracking area, and volunteers were required to carry a tag. During the data acquisition process, we manually synchronized the CRLoc and UWB systems. We split the UWB trajectory data into segments of the same length as the CSI segment to train our model.

B. CSI Preconditioning

In addition to background noise, the CSI measurement also contains other distractions owing to imperfections in the commodity WiFi NIC. Moreover, because CRLoc uses the wireless signals reflected by users to operate, the LoS signal must be removed because it is considerably stronger than the reflected signal. Next, we provide a detailed description of the preprocessing methods.

CSI measurements typically exhibit significant amplitude offsets. This leads to CSI amplitude hopping in a certain packet, and the amplitude is considerably higher than that of adjacent packets. We let the CSI amplitudes of all packets in a time window to be subtracted from the minimum amplitude. This allows the shifted CSI amplitude to return to the normal level compared to the others.



(a) Experimental device (b) Experiment in wheelhouse

(c) Experiment in hall (d) Experiment in room

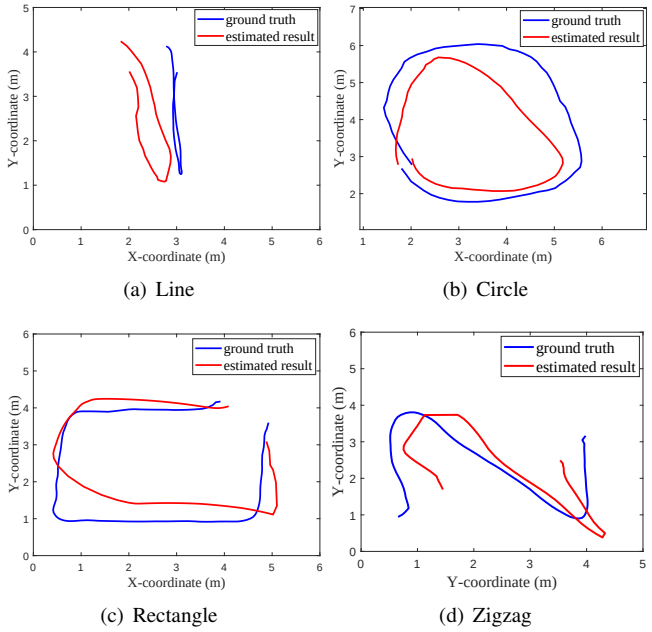
Fig. 6. Experiments device and scenes in cruise ship 'Yangtze 6'.

As for the phase error, we typically must consider three sources of error: sampling time offset (STO), carrier frequency offset (CFO), and phase-locked loop offset (PLLO). The PLLO error is constant every time the receiver starts up, and can thus be calibrated manually. The CSI phase noise caused by STO and CFO only vary in time and frequency, but not space; therefore, we select the antenna with the largest CSI mean and variance ratio as the reference antenna, and calculate the conjugate multiplications between the CSIs of each antenna and reference antenna.

Note that to eliminate the by-product term produced by conjugate multiplication, we add a constant value to the CSI amplitudes of the reference antennas, as in [32]. Finally, a bandpass Butterworth filter was used to remove the LoS signal and high-frequency noise. The cutoff frequency was set to 2 and 80 Hz caused the frequency shift owing to human movement to be in this range. The frequency shift between transceivers is zero because there is no relative movement between transceivers; therefore, the LoS signal can be removed by the bandpass Butterworth filter. We applied a time window to the CSI sequence and the length was 0.1 s, that is, we used CSIs with 0.1 s to estimate a location.

C. System Performance

We tested CRLoc and four other state-of-the-art systems in three different environments. The overall performance of CRLoc was first reported. As shown in Table. 2, CRLoc achieved a median localization error of 68 cm, 73 cm and 103 cm in a general building, ship anchored, and ship sailing



(a) Line (b) Circle

(c) Rectangle (d) Zigzag

Fig. 7. Tracking examples of different shapes of trajectories.

environment, respectively. Affected by the time-varying environment of cruise ships, the localization accuracy of CRLoc decreased slightly. But even under ship sailing conditions, 90th percentile errors were 166cm, which was much better than existing systems. Fig. 8 shows that CRLoc achieved an average localization error of 92 cm in a cruise ship environment.

We compared CRLoc with four state-of-the-art CSI-based localization techniques: Widar2.0 [13], SpotFi [20], LiFS [33] and PinLoc [34]. As shown in Table 2, these systems performed well in a general indoor environment, however, their localization errors increased rapidly on a cruise ship. This phenomenon was especially apparent when the cruise ship was sailing. The poor performances of these four systems show that they can not address the specific impact factors in a cruise ship environment. According to our analysis, the reason is two-fold. First, the dynamic characteristics of the ship make it difficult to accurately separate the target signal from the multipath signal. All these systems lack mechanisms for adapting to dynamic changes in a curise ship environment. Second, the metal material of the ship leads to high signal noise and localization confusion. These systems developed for general indoor environments fail to address the unique characteristics of cruise ships.

In addition, we studied the impact of the proposed super-resolution parameter estimation algorithm and deep learning localization model for a mobile cruise ship environment. Fig. 9 shows the effects of these two steps. The experimental results show that without the path-number estimation module, the median localization error increases to 1.31 m. This demonstrates that the algorithm can adapt to the environment of the ship and obtain more accurate channel parameters. In another test without the CRNN localization model, we diametrically located the target using the estimated parameters

Table 2 State-of-the-art system localization error: Median and 90th percentile errors tested in three different environments, that is, a general building, ship anchored and ship sailing environment.

Method		CRLoc	Widar2	Spotfi	LiFS	PinLoc
Median Error (cm)	General	68	65	51	81	97
	Anchored	73	129	84	145	186
	Sailing	103	171	148	228	263
90^{th}%ile Error (cm)	General	109	134	121	168	216
	Anchored	132	213	186	255	282
	Sailing	166	272	241	346	359

and geometric model, where the parameter with the greatest attenuation was considered as the target. In this case, the median localization error increased to 1.64 m. We believe that the CRNN model corrects the ToF–AoA estimation error and recognizes the parameters corresponding to the target.

D. Effectiveness Analysis

We discuss the effectiveness of CRLoc in terms of the following three aspects: system memory utilization, time cost, and energy consumption. We present the results from the four steps of our system: CSI preprocessing, sparse VBSAGE algorithm, ToF–AoA data generation and CRNN-based localization.

1) Memory utilization: In the preprocessing step, we need 0.35 M memory per location to store CSI and its intermediate variables. The sparse VBSAGE algorithm requires 1.2 M memory per location. After obtaining the channel parameters, we convert it into a ToF–AoA profile, which will consume 0.6 M memory. Our CRNN model requires an average of 11.2 M memory per location. In summary, the proposed system does not require significant memory.

2) Time cost: Time cost was our focus. The preprocessing and parameter estimation only take 0.038 and 0.34 s, respectively. The ToF–AoA data generation step takes 0.74 s per location with the parallel computation of four cores. Our CRNN model can obtain location coordinates within 1.2 s. Although our system does not appear to meet the real-time performance of the prototype, the time consumption of CRLoc would be considerably smaller by implementing in binary code.

3) Energy consuption: We compared the energy consumption of localization tasks and watching videos. CRLoc receives packets from AP and then calculates the user's location. According to our calculations, the amount of data traffic received by CRLoc is 69.8 k/s, which is approximately equal to the data traffic of watching videos. Therefore, the energy consumption of data transmission of CRLoc is not more than watching videos. By using the Intel extreme tuning utility, we found that the main difference is the thermal design power (TDP) of central processing unit (CPU). The TDP of localization task is 93 W, compared with 56 W for watching videos. We think the energy consumption of CRLoc is acceptable.

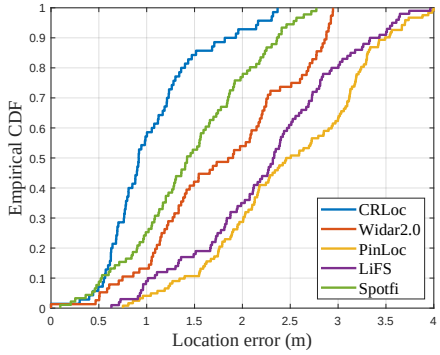


Fig. 8. Performance comparison.

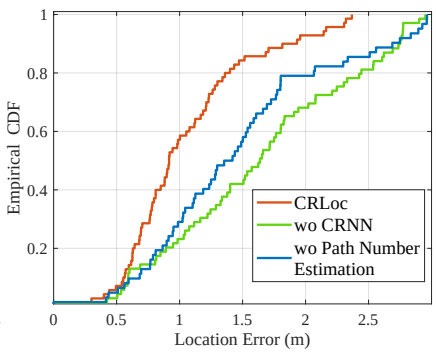


Fig. 9. Benefits of individual modules.

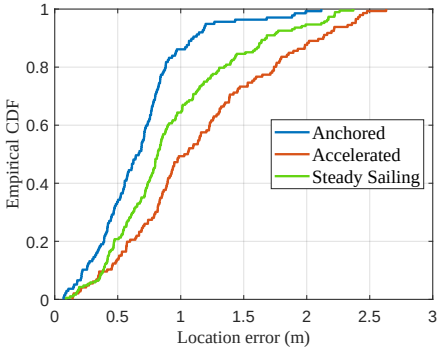


Fig. 10. Impact of ship speed state.

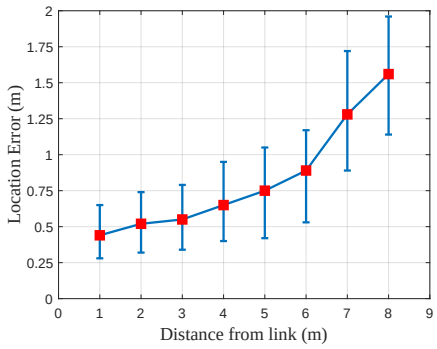


Fig. 11. Impact of walking distance from link.

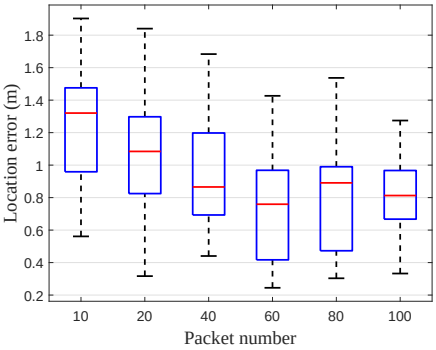


Fig. 12. Impact of packet number.

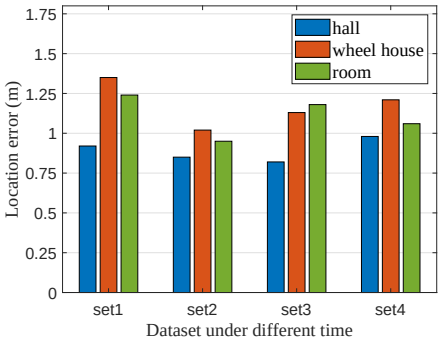


Fig. 13. Impact of different environment.

E. Parameter Disscussion

Impact of ship speed. To test the performance of our system under different ship speed states, we collected data during the ship's anchoring, acceleration, and constant-speed sailing stages. Fig. 10 shows the empirical cumulative distribution function of the tracking error of CRLoc under the three ship speed states. In the figure, we can observe that when the ship was at the anchor, the localization accuracy was the highest, reaching an error of 73 cm. When the ship started to accelerate, the location error started to increase, reaching 103 cm. When the ship entered the constant speed stage, the location error of CRLoc started to decrease, but was still greater than the error in the anchored state. This is in line with our common sense because when the ship is anchored, the hull deformation is relatively small, and the excess multipath can be eliminated by a high-pass filter. As the ship starts, the indoor environment of the ship changes drastically owing to wireless signals, which pollutes the position information carried by the wireless signals.

Impact of walking distance. To further explore the performance of the system, we studied the influence of the distance between the volunteers and Wi-Fi link. Specifically, a volunteer walked for a long distance of approximately 9 m, after which we tested our system at different distances. The evaluation was performed 20 times and the results were averaged. Fig. 11 shows the tracking error of CRLoc for targets at different distances from the Wi-Fi link. As the distance increased, the error gradually increased. This is because the farther away from the transceiver, the lower the power of the target reflection signal, which is easy to mix with metal reflection signals, such as bulkheads and furniture. The error increased rapidly after 6 m because the data we fed on the neural network were collected in a $6 \times 6m$ area.

Impact of packet number. Fig. 12 shows the effect of the number of different packages on the position error. In our experiment, we set the packet sending rate to 1000 Hz, and estimated the target localization every 0.1 s, which indicates that 100 packets were used in each estimation. We selected six different package numbers for testing. In the case of 10 packets, the overall location error was significantly larger than in the other cases. The low packet rate may lead to location aliasing because of the Doppler shift effect and at least 20 packets are required within 0.1 s to eliminate this effect [13]. However, as the number of packets increased, the location error decreased slightly. We recommend selecting a minimum of 60 packets for the location estimation.

Impact of different environment. To verify the applicability of our system in different ship areas, we selected three typical indoor ship environments. We collected data in four time periods and selected 400 locations for each environment in each period for testing. Fig. 13 shows the localization errors in three different environments. CRLoc achieved median tracking errors of 87, 113, and 104 cm for the conference hall, wheelhouse, and room of the ship, respectively. The location error in the conference hall was the

smallest because the conference hall was wide, and multipath interference was relatively small. CRLoc had a larger error in the latter two environments compared with the former because of the interference of special equipment and watchkeeper in the wheelhouse, and the rich multipath interference in the room owing to the narrow steel environment.

VIII. APPLICATIONS AND LIMITATIONS

Our study solves the passive WiFi localization problem in a cruise ship environment, which causes serious pollution in existing advanced localization systems. In this section, we introduce two practical applications of this study on cruise ships and present the limitations of our system before putting it into practice.

We primarily consider applying our work for safety and security. First, we intend to apply it to the officer of the watch (OOW) on-duty monitoring. When the ship is sailing, the OOW must perform prescribed tasks at certain locations in the wheelhouse, such as looking out, operating equipment and recording documents. Our proposed system can supervise the OOW to perform these tasks in a noninvasive manner. The second application of this study is primarily related to evacuations in emergency situations. The structure of large cruise ships is complex. When accidents such as fire and collision occur, passengers often do not know the nearest safe exit and the shortest evacuation route. Our system provides basic data for evacuation decisions, and and the evacuation system can plan the most effective evacuation route based on the location information provided.

Our system can meet certain real requirements on cruise ships, however, it has limitations that need to be addressed before it can be used in practice. Here, these limitations are described as follows:

- Our experiments were conducted in an open area within the scene, whereas some areas of the ship had more complex structures and occlusions. Under these NLoS conditions, the estimated multidimensional parameters become erroneous. Therefore, ensuring localization accuracy under NLoS conditions is our future study.
- The size of the localization area in this study was limited to 6m$\times$6m, and the localization accuracy reduced significantly when the target was outside this area. Enabling CRLoc to run in any area is an improvement direction.
- The model established in this study is only for the single-person localization problem, whereas multi-person localization is more in line with real needs. Therefore, this is also a limitation of CRLoc and motivates for future studies.

IX. CONCLUSION

In this paper, we designed and implemented one of the first WiFi-based passive motion tracking system in a cruise ship environment. We conducted extensive experiments to identify the main factors influencing localization in a mobile environment and to study the relationship between CSI variation and environmental dynamics. We utilize a super-resolution parameter estimation algorithm to combat dynamic multipaths in a cruise ship. Then, we designed a deep learning-based approach as the tracking model, which improved the accuracy of location estimation. Real-world implementation and evaluation show that CRLoc achieves a median location accuracy of 0.92 m in a 6 m $\times$ 6 m area and can meet the localization and tracking requirements on a cruise ship.

ACKNOWLEDGMENT

This study was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 51979216, the Natural Science Foundation of Hubei Province, China, under Grant No.2021CFA001.

REFERENCES

[1] K. Mizumoto, K. Kagaya, A. Zarebski, and G. Chowell, "Estimating the asymptomatic proportion of coronavirus disease 2019 (covid-19) cases on board the diamond princess cruise ship, yokohama, japan, 2020," *Eurosurveillance*, vol. 25, no. 10, p. 2000180, 2020.

[2] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[3] J. Xu, H. Chen, K. Qian, E. Dong, M. Sun, C. Wu, L. Zhang, and Z. Yang, "ivr: Integrated vision and radio localization with zero human effort," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–22, 2019.

[4] Y. Zhao, J. Xu, J. Wu, J. Hao, and H. Qian, "Enhancing camera-based multimodal indoor localization with device-free movement measurement using wifi," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1024–1038, 2019.

[5] T. Ma, Y. Xiao, X. Lei, W. Xiong, and Y. Ding, "Indoor localization with reconfigurable intelligent surface," *IEEE Communications Letters*, vol. 25, no. 1, pp. 161–165, 2020.

[6] T. Liu, B. Li, and L. Yang, "Phase center offset calibration and multi-point time latency determination for uwb location," *IEEE Internet of Things Journal*, 2022.

[7] Y. Ma, N. Selby, and F. Adib, "Minding the billions: Ultra-wideband localization for deployed rfid tags," in *Proceedings of the 23rd annual international conference on mobile computing and networking*, 2017, pp. 248–260.

[8] X. Liu, J. Zhang, S. Jiang, Y. Yang, K. Li, J. Cao, and J. Liu, "Accurate localization of tagged objects using mobile rfid-augmented robots," *IEEE Transactions on Mobile Computing*, 2019.

[9] K. Ngamakeur, S. Yongchareon, J. Yu, and S. U. Rehman, "A survey on device-free indoor localization and tracking in the multi-resident environment," *ACM Computing Surveys (CSUR)*, vol. 53, no. 4, pp. 1–29, 2020.

[10] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your csi: A channel state information extraction platform for modern wi-fi chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, 2019, pp. 21–28.

[11] M. Chen, K. Liu, J. Ma, Y. Gu, Z. Dong, and C. Liu, "Swim: Speed-aware wifi-based passive indoor localization for mobile ship environment," *IEEE Transactions on Mobile Computing*, 2019.

[12] M. Chen, K. Liu, J. Ma, X. Zeng, Z. Dong, G. Tong, and C. Liu, "Moloc: Unsupervised fingerprint roaming for device-free indoor localization in a mobile ship environment," *IEEE Internet of Things Journal*, vol. 7, no. 12, pp. 11 851–11 862, 2020.

[13] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2. 0: Passive human tracking with a single wi-fi link," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, 2018, pp. 350–361.

[14] X. Sun, H. Ai, J. Tao, T. Hu, and Y. Cheng, "Bert-adloc: A secure crowdsourced indoor localization system based on ble fingerprints," *Applied Soft Computing*, vol. 104, p. 107237, 2021.

[15] Y. Yu, R. Chen, L. Chen, X. Zheng, D. Wu, W. Li, and Y. Wu, "A novel 3-d indoor localization algorithm based on ble and multiple sensors," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9359–9372, 2021.

[16] X. Chen, C. Ma, M. Allegue, and X. Liu, "Taming the inconsistency of wi-fi fingerprints for device-free passive indoor localization," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE, 2017, pp. 1–9.

[17] X. Chen, H. Li, C. Zhou, X. Liu, D. Wu, and G. Dudek, "Fidora: Robust wifi-based indoor localization via unsupervised domain adaptation," *IEEE Internet of Things Journal*, 2022.

[18] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single wifi access point," in *13th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 16)*, 2016, pp. 165–178.

[19] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-music: accurate device-free indoor localization," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 196–207.

[20] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using wifi," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 2015, pp. 269–282.

[21] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "md-track: Leveraging multi-dimensionality for passive indoor wi-fi tracking," in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–16.

[22] Z. Chen, G. Zhu, S. Wang, Y. Xu, J. Xiong, J. Zhao, J. Luo, and X. Wang, "m^3: Multipath assisted wi-fi localization with a single access point," *IEEE Transactions on Mobile Computing*, 2019.

[23] R. H. Venkatnarayan, M. Shahzad, S. Yun, C. Vlachou, and K.-H. Kim, "Leveraging polarization of wifi signals to simultaneously track multiple people," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 2, pp. 1–24, 2020.

[24] D. Shutin and B. H. Fleury, "Sparse variational bayesian sage algorithm with application to the estimation of multipath wireless channels," *IEEE Transactions on signal processing*, vol. 59, no. 8, pp. 3609–3623, 2011.

[25] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Zero-effort cross-domain gesture recognition with wi-fi," in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, 2019, pp. 313–325.

[26] W. Mao, M. Wang, W. Sun, L. Qiu, S. Pradhan, and Y.-C. Chen, "Rnn-based room scale hand motion tracking," in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–16.

[27] M. Lundgren, L. Svensson, and L. Hammarstrand, "Variational bayesian expectation maximization for radar map estimation," *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1391–1404, 2015.

[28] B. Sun, Y. Guo, N. Li, and D. Fang, "Multiple target counting and localization using variational bayesian em algorithm in wireless sensor networks," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2985–2998, 2017.

[29] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 11, pp. 2298–2304, 2016.

[30] L. Perotin, R. Serizel, E. Vincent, and A. Guerin, "Crnn-based multiple doa estimation using acoustic intensity features for ambisonics recordings," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 1, pp. 22–33, 2019.

[31] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 4, pp. 159–170, 2010.

[32] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, "Indotrack: Device-free indoor human tracking with commodity wi-fi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–22, 2017.

[33] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie, "Lifs: Low human-effort, device-free localization with fine-grained subcarrier information," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, 2016, pp. 243–256.

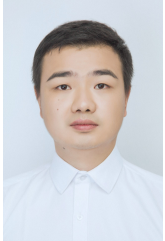
[34] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka, "You are facing the mona lisa: Spot localization using phy layer information," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*, 2012, pp. 183–196.

Kezhong Liu received the B.S. and M.S. degrees in marine navigation from the Wuhan University of Technology(WUT), Wuhan, China, in 1998 and 2001, respectively. He received the Ph.D. degree in communication and information engineering from the Huazhong University of Science and Techonogy, Wuhan, China, in 2006. He is currently a professor with School of Navigation, WUT. His active research interests include indoor localization technology and data mining for ship navigation.

Wen Yang received the B.S. degree in mechatronic engineering from the Guizhou university, Guizhou, China, in 2018. He is currently a master student in the Wuhan University of Technology (WUT). His research interests include wireless localization and sensing.

Mozi Chen received the B.S. degree in electric engineering from the Hubei University of Technology, China, in 2013. He received the M.S. and Ph.D.degree in navigation engineering from the Wuhan University of Technology (WUT), China, in 2016 and 2020. He is currently an associate researcher in WUT. His research work has been focusing on wireless sensing techniques and machine learning algorithms for human localization, emergency navigation and activity recognition in mobile environment, i.e., cruise ships.

Kai Zheng received the Ph.D. degree with the School of Geodesy and Geomatics, Wuhan University, in 2020. He is currently an associate researcher with the Wuhan University of Technology. His research interests are GNSS precise positioning techniques.



Xuming Zen received the Ph.D. degree in communication engineering from the China University of Geosciences, Wuhan, China, in 2018. He is currently pursuing the Ph.D. degree in traffic and transportation engineering with the School of Navigation, Wuhan University of Technology, Wuhan. He has been studying wireless ad-hoc network for shipboard environment with the Wuhan University of Technology since 2019. His research interests include routing protocols, MAC, QoS, clustering, and reliable wireless transmission.

Shengkai Zhan received his Ph.D. degree from the EIC Department, Huazhong University of Science and Technology (HUST) in 2021. He received his M.Sc and M.Phil degrees from HUST and Hong Kong University of Science and Technology in 2012 and 2014, respectively. He is currently an associate professor at Wuhan University of Technology (WUT). His recent research interests include state estimation, wireless sensing, mobile computing, multi-sensor fusion, robot control and planning.

Cong Liu received the Ph.D. degree in computer science from the University of North Carolina at Chapel Hill, in Jul. 2013. He is currently an associate professor in the the Department of Electrical and Computer Engineering, University of California, Riverside. His research interests include realtime systems and GPGPU. He has published more than 30 papers in premier conferences and journals. He received the Best Paper Award at the 30th IEEE RTSS and the 17th RTCSA. He is a member of the IEEE.