



UNIVERSITY OF LEEDS

This is a repository copy of *Online Spatiotemporal Modeling for Robust and Lightweight Device-Free Localization in Nonstationary Environments*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/192729/>

Version: Accepted Version

Article:

Zhang, J, Li, Y, Xiao, W et al. (1 more author) (2022) Online Spatiotemporal Modeling for Robust and Lightweight Device-Free Localization in Nonstationary Environments. IEEE Transactions on Industrial Informatics. ISSN 1551-3203

<https://doi.org/10.1109/TII.2022.3218666>

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Online Spatiotemporal Modeling for Robust and Lightweight Device-Free Localization in Nonstationary Environments

Jie Zhang, *Member, IEEE*, Yanjiao Li, *Member, IEEE*, Wendong Xiao, *Senior Member, IEEE*, and Zhiqiang Zhang, *Member, IEEE*

Abstract—Recent advances in WiFi-based device-free localization (DFL) mainly focus on stationary scenarios, and ignore the environmental dynamics, hindering the large-scale implementation of the DFL technique. In order to enhance the localization performance in nonstationary environments, in this paper, a novel multidomain collaborative extreme learning machine (MC-ELM)-based DFL framework is proposed. Specifically, the whole environment is first divided into several sub-domains depending on the distributions of the collected data using clustering algorithm, and a corresponding number of local DFL models are then built to represent these sub-domains separately. Finally, a global DFL model is achieved through seamlessly integrating all the local DFL models in a global optimization manner. The created MC-ELM-based DFL model also can be incrementally updated with sequentially coming data without retraining to track the environmental dynamics. Extensive experiments in several indoor environments demonstrate the robustness and generalization of the proposed MC-ELM-based DFL framework.

Index Terms—Device-free localization, multidomain representation, nonstationary environments, model incremental updating.

I. INTRODUCTION

DUE to the rapid development of wireless communication techniques and ubiquitous deployment of Internet of Things (IoT) infrastructures, location-based services are powering and reshaping our daily life [1], [2], [3]. Prior studies have shown that the existence of the target posing influence on surrounding wireless signals can be utilized to estimate target's location [4], [5], [6]. Sequentially, WiFi-based device-free localization (DFL) techniques were proposed, and have

This work was supported in part by the National Natural Science Foundation of China under Grant 62003038, and in part by the Fundamental Research Fund of University of Science and Technology Beijing under Grant 00007729.

Jie Zhang is with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China, also with the Beijing Engineering Research Center of Industrial Spectrum Imaging, Beijing 100083, China, and also with the School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, U.K. (e-mail: zhangjie_sae@ustb.edu.cn)

Yanjiao Li is with the Institute of Engineering Technology, University of Science and Technology Beijing, Beijing 100083, China, and also with the National Engineering Research Center for Advanced Rolling and Intelligent Manufacturing, University of Science and Technology Beijing, Beijing 100083, China (e-mail: yanjiaoli@ustb.edu.cn)

Wendong Xiao is with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China, and also with the Beijing Engineering Research Center of Industrial Spectrum Imaging, Beijing 100083, China (e-mail: wdxiao@ustb.edu.cn)

Zhiqiang Zhang is with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K., LS2 9JT (e-mail: z.zhang3@leeds.ac.uk).

attracted great attention in both academia and industry, due to the ubiquitous, low-cost, and non-intrusive characteristics of WiFi signals [7], [8], [9]. Different from device-based and computer vision-based localization techniques, WiFi-based DFL does not need the target to equip with dedicated electronic devices and the target's location can be estimated passively, and has enough sensing coverage and privacy protection capability without lighting requirements, enabling it a promising candidate technique for several intelligent IoT applications [10], [11], [12], [13]. Early WiFi-based DFL methods mainly utilize received signal strength (RSS), but the localization performance of such set of methods is seriously limited due to the unpredictable fluctuations of RSS. In recent years, channel state information (CSI) can be retrieved from network interface cards (NIC) with CSTOOL [14], and CSI has the better characterization of the frequency response of the wireless channel at orthogonal frequency-division multiplexing (OFDM) sub-carriers, which can enhance the localization performance of WiFi-based DFL.

In the past few years, great progresses have been made in WiFi-based DFL, but there are still two major issues hindering the large-scale implementation of the DFL technique, including: 1) vulnerable to environmental dynamics, and 2) high human effort cost for DFL model retraining. Accurate localization highly relies on building a good DFL model using massive collected data with consistent pair-to-pair feature mappings. However, pattern consistency between signal variations caused by the existence of the target and the corresponding locations cannot hold due to different types of interference, such as opening and closing windows or doors, layout changes of environments, and thermal drifts of WiFi devices, etc. When the radio map created for one specific environment at a specific time step is utilized to estimate the target's location of other time steps, the real-time signal measurements may seriously deviate from the fingerprints in the created radio map during the online localization phase due to the environmental dynamics.

Considering that transfer learning has the potential to transfer knowledge from a well-labeled domain (i.e., source domain) to an unlabeled domain (i.e., target domain), state-of-the-art methods adopt transfer learning to accommodate environmental dynamics. Specifically, Wang et al. [15] designed a modified DFL scheme to process WiFi measurements, which not only could eliminate the need of acquiring reference WiFi measurements, but also alleviated the negative effects incurred

by the surrounding environments. Ohara et al. [16] transferred a signal strength model from the source domain to the target domain, but the proposed method required the link lengths of different environments to be the same, which might not hold in practical scenarios. Chang et al. [17] designed FitLoc for multiple targets' location estimation, which could transfer the radio map between the source environment and the target environment, but the construction of the radio map of the source environment was time-consuming. DTransfer [18] was designed, in which an optimized low-rank matrix completion model was proposed for efficiently constructing the radio map. Chen et al. [19] proposed a LSTM-based transfer model for mobile ship environment, due to the immediate, dynamic and unpredictable changes of the ship during voyages, to enhance the localization performance. In addition, transferring compressive-sensing-based DFL method for different categories of targets was also investigated [20]. The essence of transfer learning is to find an equivalent transformation to minimize the distance between the source domain and the target domain. However, environmental dynamics may enlarge the difference between the two domains in nonstationary scenarios. Transfer learning-based methods need to retrain the existing DFL model frequently to mitigate the gap between the source domain and the target domain, which is time-consuming and labor-intensive. It is worthwhile to develop advanced DFL method to improve the localization performance and practicability in real indoor settings by simultaneously strengthening robustness and adaptability of environmental dynamics, and reducing human effort on model retraining.

In order to tackle the aforementioned issues, in this paper, we propose a multidomain collaborative extreme learning machine (MC-ELM)-based DFL framework. In terms of robustness, MC-ELM-based DFL framework fully considers the spatiotemporal information of the cluttered environment, and builds a global DFL model through global optimization. In terms of lightweight modeling, MC-ELM-based DFL framework utilizes the newly collected data to update the existing DFL model incrementally without retraining to adapt to the new environment, significantly reducing the human effort cost in dealing with environmental dynamics. The proposed MC-ELM-based DFL framework involves two phases, including the model initialization phase and the model incremental updating phase. Specifically, the whole environment is first divided into several signal-level sub-domains, where the CSI signals with similar statistical characteristics will be clustered into the same sub-domains. In this manner, the distance among the CSI signals in a cluster are smaller, while the distance among clusters are larger. In the model initialization phase, a corresponding number of local DFL models are built using ELMs for separately representing all the sub-domains. After that, an initial global DFL model is built through seamlessly integrating all the local DFL models. In the model incremental updating phase, when the newly collected data of other time steps are available, the created global DFL model could be incrementally updated without retraining, enabling it to track various environmental dynamics. Different from state-of-the-art methods, the proposed MC-ELM-based DFL framework fully considers the spatial properties of the physical space and

time dynamics of the environment, making it more robust in nonstationary scenarios. The main contributions can be summarized as follows:

- 1) A local-global modeling framework for robust DFL in nonstationary scenarios is developed to characterize the environment separately using the local DFL models, and then all the local DFL models are seamlessly integrated into a global DFL model. In this manner, the whole environment could be represented better by the local DFL models, and the target's location could be estimated by the global DFL model accurately and robustly.
- 2) A modified lightweight distributed online updating mechanism is proposed to cope with environmental dynamics. When the data newly collected from other time steps are available, the existing DFL model could be updated through the proposed distributed online updating mechanism in a lightweight manner without retraining.
- 3) We implement the proposed framework with commodity WiFi devices and conduct comprehensive experiments in real indoor settings. Experimental results demonstrate that it could achieve better localization performance both in stationary and nonstationary scenarios, indicating its robustness and great generalization performance.

The remaining of this paper is organized as follows: Basics of CSI and problem formulation are presented in Section II. Section III details the proposed method. Performance evaluation and further analysis are reported in Section IV. Some issues, limitations and future directions are discussed in Section V. Finally, conclusions are given in Section VI.

II. PRELIMINARIES AND PROBLEM FORMULATION

In this section, we first introduce the basics of CSI, and then analyze the effects of environmental dynamics on DFL to motivate the problem formulation.

A. CSI Primer

Along with OFDM, multiple-input multiple-out (MIMO) could provide CSI of each subcarrier with both amplitude and phase information. CSI could describe how the wireless signals propagate through the physical space in time, frequency, and spatial domains. In a typical indoor environment, WiFi signals propagate from the transmitter to the receiver via multiple paths, including one direct path and several paths from ambient objects, such as furniture, ceiling, wall, and floor, etc. CSI is actually a superposition of signals from all the paths. When a person enters in the monitoring area, additional signal paths are introduced by the scattering of human body, and the CSI measurements could characterize the effects of human presence. Thus, the location can be estimated using the collected CSI measurements.

B. Problem Formulation

As mentioned above, state-of-the-art methods mainly focus on the target's location estimation in the stationary environment at a specific time step, assuming that the collected

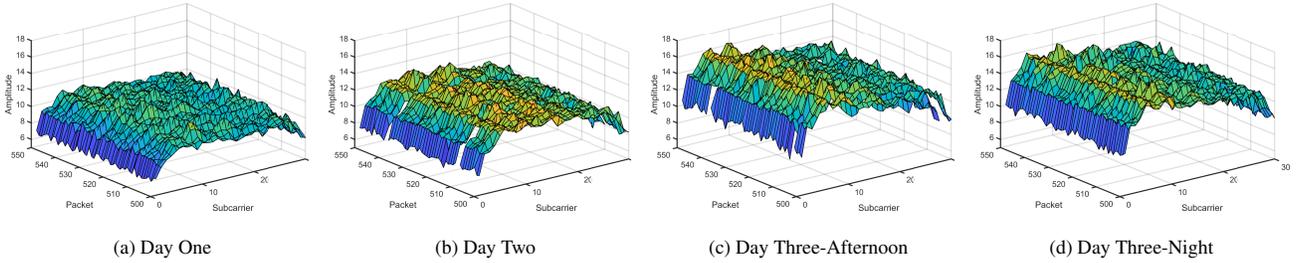


Fig. 1. Empirical Study of Environmental Dynamics

data are subject to uniform distribution. However, such kind of assumption usually cannot hold in real-world application scenarios, due to the seasonality or periodicity effects (such as opening/closing window/door), function faults of the sensors embedded in WiFi devices, thermal drifts and/or aging effects of WiFi devices [21]. Moreover, WiFi signals transmit to the receiver not only via the Line-of-Sight (LoS) path, but also several reflection paths due to the multipath propagation in the cluttered environment. Several signal propagation paths will be changed once the layout of the environment varies, and the statistical characteristics of the CSI measurements will also change, making the localization performance of the existing DFL model degrade [7]. In this section, we analyze the effects of environmental dynamics on data distribution from different aspects, which motivate the problem formulation.

Mathematically, for a specific environment, let $(\mathbf{x}^t, \mathbf{y}^t)$ be the data pair sampled from an unknown probability distribution $P(\mathbf{x}^t, \mathbf{y}^t)$ at time step t , $P(\mathbf{y}^t|\mathbf{x}^t)$ and $P(\mathbf{x}^t)$ be the corresponding posterior distribution and evidence distribution, respectively. When the data sequentially come chunk-by-chunk, $P(\mathbf{y}^t|\mathbf{x}^t)$ may vary over time, independently from variations in $P(\mathbf{x}^t)$. This is not consistent with the assumption that data utilized for building DFL model follow the uniform distribution, easily leading to performance loss.

Experimentally, Fig. 1 depicts the CSI amplitudes of consecutive three days and different time segments in one specific day without target in the environment. According to Fig. 1 (a-c), we can find that CSI amplitudes of different days are obviously different. Additionally, even in the same day, CSI amplitudes are different in local details (see Fig. 1 (c) and Fig. 1 (d)). This empirical study verifies the above mathematical analysis, indicating the fact that one of the effective ways to alleviate the negative effects of environmental dynamics on the localization performance is to frequently and periodically update the existing DFL model using the sequentially collected data.

In summary, both mathematical analysis and empirical study demonstrate that the traditional “non-adaptive” DFL methods under the stationary assumption cannot function well in nonstationary scenarios, due to the time-varying distribution properties of the data. Thus, it is worthwhile to develop modified DFL method to enhance the localization performance in nonstationary environments.

III. METHODOLOGY

In this section, we detail the main framework of MC-ELM-based DFL, local DFL models for sub-domains, MC-ELM-based global DFL model, and the corresponding model initialization and incremental updating strategy, respectively.

A. Main Framework of MC-ELM-based DFL

In order to perform robust and lightweight DFL in non-stationary environments, we propose a modified MC-ELM-based DFL method, which can fully consider both the spatial properties of the physical space and time dynamics of the environment. Specifically, as illustrated in Fig. 2, the proposed method involves two phases, including the model initialization phase and the model incremental updating phase. The collected data are first divided into several groups by clustering algorithm, making the data with uniform distributions be in the same groups. In the model initialization phase, a corresponding number of local DFL models are built based on ELM with those grouped data. Additionally, all the local DFL models are seamlessly integrated into a global DFL model and optimized through a global optimization manner. In the model incremental updating phase, the created global DFL model could be updated recursively without retraining using the newly coming data collected from other time steps. In this manner, the existing DFL model could be updated fast and frequently, making it keep satisfactory localization performance under environmental dynamics.

B. Local DFL Models for Sub-domains

State-of-the-art methods usually utilize the whole collected data to build a general DFL model under the assumption that all those data are subject to uniform distribution. However, data collected from a multipath-rich environment may follow different mathematical distributions. Thus, we build several local DFL models for different sub-domains to separately represent them. To be specific, the dataset is first divided into v clusters through K-means clustering algorithm. Each cluster is represented as $\{(\mathbf{x}_{pq}, \mathbf{y}_{pq})\}_{p=1,2,\dots,v,q=1,2,\dots,\tilde{N}_p}$, where \mathbf{x}_{pq} and \mathbf{y}_{pq} are the q th input and output in the p th cluster, and \tilde{N}_p denotes the number of data in the corresponding cluster. After that, ELM is employed to represent each sub-domain:

$$f_p(\mathbf{x}_p) = \mathbf{s}(\mathbf{x}_p)\boldsymbol{\beta}_p \quad (1)$$

where $\mathbf{s}(\mathbf{x}_p)$ is the hidden layer output matrix of the p th cluster, for mapping the data from the input space to the

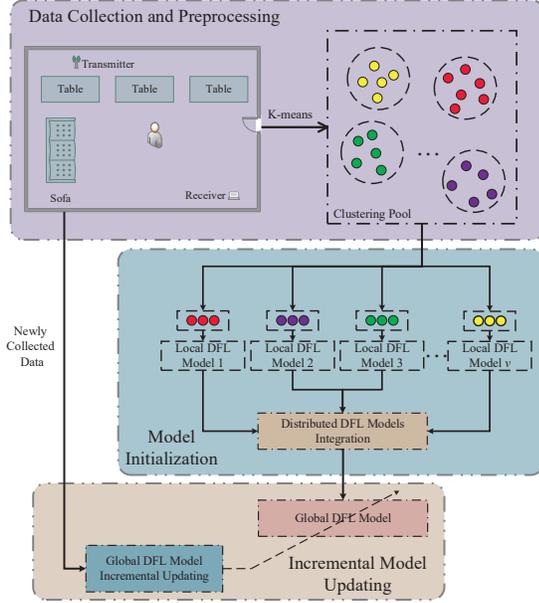


Fig. 2. Multidomain Representation-based Online Modeling for DFL

feature space, and β_p denotes the output weights of the p th local DFL model. Different from conventional machine learning algorithms, hidden layer parameters of ELM in $s(\mathbf{x}_p)$ are randomly generated before training, reducing the learning problem to that of identifying the optimal output weights β_p . Therefore, ELM has faster learning speed and better generalization performance.

According to [22], the objective function of ELM is to simultaneously minimize the norm of output weights and the training error:

$$\begin{aligned} \min : & \frac{1}{2} \|\beta_p\|^2 + \frac{1}{2} \lambda \sum_{q=1}^{\tilde{N}_p} \varsigma_{pq}^2 \\ \text{s.t.}, & s(\mathbf{x}_{pq}) \beta_p = \mathbf{y}_{pq} - \varsigma_{pq}, p = 1, \dots, v, q = 1, \dots, \tilde{N}_p \end{aligned} \quad (2)$$

where λ represents the regularization coefficient, and ς_{pq} represents the training error, respectively.

C. MC-ELM-based Global DFL Model

There are intersections between the adjacent clusters, the main reason is that different local layouts (such as surrounding environments and deployments of WiFi devices, etc.) lead to the different distributions of the corresponding collected data, but such kind of difference is not obvious at the edge of the adjacent clusters. Thus, the global DFL model should first consider the rich multipath caused by the layout of the environment, and then balance the global localization performance and the relevance of the divided sub-domains:

$$\begin{aligned} \min : & \frac{1}{2} \sum_{p=1}^v \|\beta_p\|^2 + \frac{1}{2} \lambda \sum_{p=1}^v \sum_{q=1}^{\tilde{N}_p} \varsigma_{pq}^2 + \\ & \frac{1}{2} \mu \sum_{p=1}^v \left\| \beta_p - \frac{1}{v-1} \sum_{l=1, l \neq p}^v \beta_l \right\|^2 \\ \text{s.t.}, & s(\mathbf{x}_{pq}) \beta_p = \mathbf{y}_{pq} - \varsigma_{pq}, p = 1, \dots, v, q = 1, \dots, \tilde{N}_p \end{aligned} \quad (3)$$

where μ represents the regulation coefficient for reducing the gaps between local DFL models, β_p and β_l represent the output weights of local DFL models, respectively.

The salient features of the modified objective function can be summarized as follows:

- 1) The first term is to minimize the output weights of local DFL models to strengthen the generalization performance of the global DFL model, and also avoid the over-fitting problem during the model training phase.
- 2) The second term is to minimize the training errors of local DFL models, which could enhance the localization performance. A good DFL model should have small training error.
- 3) The third term is to minimize the difference of the output weights between local DFL models. It considers the interaction, continuity, and smoothness between local DFL models, and reduces the gaps between local DFL models. This term makes all the local DFL models more collaboratively, and could improve the global generalization performance.

D. Global Model Initialization

In the model initialization phase, an initial DFL model is created to represent the current environment using the data collected from the current time step t_0 , denoted as \mathbb{D}_0 , which has \tilde{N}^0 samples.

In order to solve the objective function demonstrated in (3), we should first perform a transformation (The detailed transformation process is shown in Appendix). The equivalent objective function of MC-ELM is

$$\begin{aligned} \min : & \frac{1}{2} \|\bar{\beta}^{(0)}\|^2 + \frac{1}{2} \lambda \sum_{p=1}^v \sum_{q=1}^{\tilde{N}_p^0} (\varsigma_{pq}^0)^2 \\ \text{s.t.}, & \bar{s}_0(\mathbf{x}_{pq}^0) \bar{\beta}^{(0)} = \mathbf{y}_{pq}^0 - \varsigma_{pq}^0, p = 1, \dots, v, q = 1, \dots, \tilde{N}_p^0 \end{aligned} \quad (4)$$

where $\bar{s}_0(\mathbf{x}_{pq}^0)$ is the corresponding hidden layer output matrix of the new equivalent output weights $\bar{\beta}^{(0)}$.

According to the Karush-Kuhn-Tucker theorem, the dual optimization problem of (4) can be expressed by (5). In (5), α_{pq} is the Lagrange multiplier corresponding to the q th training sample in the p th local DFL model, α^p , \bar{s}_0^p , \mathbf{y}_0^p and ς_0^p represent the Lagrange multiplier, hidden layer output matrix, ground truth, and training error of the p th local DFL model, respectively. Therefore, the initial output weights of MC-ELM can be calculated by (6), and then the initial DFL model is established (both (5) and (6) are shown at the top of the next page).

E. Global Model Incremental Updating

As mentioned above, various environmental dynamics could vary the statistical characteristics of the data collected at different time steps. It means that if we still use the DFL model created in the initialization phase to estimate the target's location of other time steps, the localization performance may be seriously degraded. Thus, we should update the existing DFL model with the newly collected data, making the DFL model can track the environmental dynamics.

$$\begin{aligned}
L(\bar{\beta}^{(0)}, \varsigma_{pq}^0, \alpha_{pq}) &= \frac{1}{2} \|\bar{\beta}^{(0)}\|^2 + \frac{1}{2} \lambda \sum_{p=1}^v \sum_{q=1}^{\bar{N}_p^0} (\varsigma_{pq}^0)^2 - \sum_{p=1}^v \sum_{q=1}^{\bar{N}_p^0} \alpha_{pq} \left(\bar{s}_0(\mathbf{x}_{pq}^0) \bar{\beta}^{(0)} - \mathbf{y}_{pq}^0 + \varsigma_{pq}^0 \right) \\
&= \frac{1}{2} \|\bar{\beta}^{(0)}\|^2 + \frac{1}{2} \lambda \sum_{p=1}^v \|\varsigma_p^0\|^2 - [\alpha^1 \alpha^2 \dots \alpha^v] \left(\begin{bmatrix} \bar{s}_0^1 \\ \bar{s}_0^2 \\ \vdots \\ \bar{s}_0^v \end{bmatrix} \bar{\beta}^{(0)} - \begin{bmatrix} \mathbf{y}_0^1 \\ \mathbf{y}_0^2 \\ \vdots \\ \mathbf{y}_0^v \end{bmatrix} + \begin{bmatrix} \varsigma_0^1 \\ \varsigma_0^2 \\ \vdots \\ \varsigma_0^v \end{bmatrix} \right)
\end{aligned} \tag{5}$$

$$\bar{\beta}^{(0)} = \left[(\bar{s}_0^1)^T \bar{s}_0^1 + (\bar{s}_0^2)^T \bar{s}_0^2 + \dots + (\bar{s}_0^v)^T \bar{s}_0^v + \frac{\mathbf{I}}{\lambda} \right]^{-1} \begin{bmatrix} (\bar{s}_0^1)^T & (\bar{s}_0^2)^T & \dots & (\bar{s}_0^v)^T \end{bmatrix} \begin{bmatrix} \mathbf{y}_0^1 \\ \mathbf{y}_0^2 \\ \vdots \\ \mathbf{y}_0^v \end{bmatrix} \tag{6}$$

In the model incremental updating phase, when the data collected from the next time step t_1 are available, they can be employed to obtain the new output weights $\bar{\beta}^{(1)}$, and then the existing global DFL model could be updated. Let \mathbb{D}_1 be the corresponding data, and \bar{N}^1 denotes the number of available training data at t_1 .

In order to simplify the following derivation, let

$$Q_0 = \left(\bar{\mathbf{S}}_0^T \bar{\mathbf{S}}_0 + \frac{\mathbf{I}}{\lambda} \right)^{-1} \tag{7}$$

where $\bar{\mathbf{S}}_0 = \begin{bmatrix} \bar{s}_0^1 \\ \bar{s}_0^2 \\ \vdots \\ \bar{s}_0^v \end{bmatrix}$.

The hidden layer output matrix of the time step t_1 is

$$\bar{\mathbf{S}}_1 = \begin{bmatrix} \bar{s}_1^1 \\ \bar{s}_1^2 \\ \vdots \\ \bar{s}_1^v \end{bmatrix} \tag{8}$$

Thus, we have

$$Q_1 = \left((\bar{s}_1^1)^T \bar{s}_1^1 + (\bar{s}_1^2)^T \bar{s}_1^2 + \dots + (\bar{s}_1^v)^T \bar{s}_1^v + \frac{\mathbf{I}}{\lambda} \right)^{-1} \tag{9}$$

where $\bar{s}_1^1 = \begin{bmatrix} \bar{s}_0^1 \\ \bar{s}_1^1 \end{bmatrix}$, $\bar{s}_1^2 = \begin{bmatrix} \bar{s}_0^2 \\ \bar{s}_1^2 \end{bmatrix}$, ..., $\bar{s}_1^v = \begin{bmatrix} \bar{s}_0^v \\ \bar{s}_1^v \end{bmatrix}$. In addition, $\bar{s}_1^1, \bar{s}_1^2, \dots, \bar{s}_1^v$ denote the hidden layer output matrices of the corresponding new data chunks.

For the following model updating, we expand (9):

$$\begin{cases} (\bar{s}_1^1)^T \bar{s}_1^1 = (\bar{s}_0^1)^T \bar{s}_0^1 + (\bar{s}_1^1)^T \bar{s}_1^1 \\ (\bar{s}_1^2)^T \bar{s}_1^2 = (\bar{s}_0^2)^T \bar{s}_0^2 + (\bar{s}_1^2)^T \bar{s}_1^2 \\ \vdots \\ (\bar{s}_1^v)^T \bar{s}_1^v = (\bar{s}_0^v)^T \bar{s}_0^v + (\bar{s}_1^v)^T \bar{s}_1^v \end{cases} \tag{10}$$

Accordingly, we have

$$\begin{aligned}
Q_1^{-1} &= \left((\bar{s}_0^1)^T \bar{s}_0^1 + (\bar{s}_1^1)^T \bar{s}_1^1 \right) + \left((\bar{s}_0^2)^T \bar{s}_0^2 + (\bar{s}_1^2)^T \bar{s}_1^2 \right) + \dots + \left((\bar{s}_0^v)^T \bar{s}_0^v + (\bar{s}_1^v)^T \bar{s}_1^v \right) + \frac{\mathbf{I}}{\lambda} \\
&= Q_0^{-1} + (\bar{s}_1^1)^T \bar{s}_1^1 + (\bar{s}_1^2)^T \bar{s}_1^2 + \dots + (\bar{s}_1^v)^T \bar{s}_1^v
\end{aligned} \tag{11}$$

Multiply Q_0 at both sides in (11):

$$Q_0 Q_1^{-1} = \mathbf{I} + Q_0 (\bar{s}_1^1)^T \bar{s}_1^1 + Q_0 (\bar{s}_1^2)^T \bar{s}_1^2 + \dots + Q_0 (\bar{s}_1^v)^T \bar{s}_1^v \tag{12}$$

Multiply Q_1 at both sides, (12) could be modified as

$$Q_0 = Q_1 + Q_0 (\bar{s}_1^1)^T \bar{s}_1^1 Q_1 + Q_0 (\bar{s}_1^2)^T \bar{s}_1^2 Q_1 + \dots + Q_0 (\bar{s}_1^v)^T \bar{s}_1^v Q_1 \tag{13}$$

Then, we obtain

$$Q_1 = \left(\mathbf{I} + Q_0 (\bar{s}_1^1)^T \bar{s}_1^1 + Q_0 (\bar{s}_1^2)^T \bar{s}_1^2 + \dots + Q_0 (\bar{s}_1^v)^T \bar{s}_1^v \right)^{-1} Q_0 \tag{14}$$

The initial output weights of MC-ELM could be calculated:

$$\begin{aligned}
\bar{\beta}^{(0)} &= Q_0 \bar{\mathbf{S}}_0^T \mathbf{Y}_0 \\
&= Q_0 \left((\bar{s}_0^1)^T \mathbf{y}_0^1 + (\bar{s}_0^2)^T \mathbf{y}_0^2 + \dots + (\bar{s}_0^v)^T \mathbf{y}_0^v \right)
\end{aligned} \tag{15}$$

where $\mathbf{Y}_0 = \begin{bmatrix} \mathbf{y}_0^1 \\ \mathbf{y}_0^2 \\ \vdots \\ \mathbf{y}_0^v \end{bmatrix}$.

Therefore, we have the following expression of the new output weights of time step t_1 :

$$\bar{\beta}^{(1)} = Q_1 \left((\bar{s}_1^1)^T \mathbf{y}_1^1 + (\bar{s}_1^2)^T \mathbf{y}_1^2 + \dots + (\bar{s}_1^v)^T \mathbf{y}_1^v \right) \tag{16}$$

where $\mathbf{y}_1^1 = \begin{bmatrix} \mathbf{y}_0^1 \\ \eta_1^1 \end{bmatrix}$, $\mathbf{y}_1^2 = \begin{bmatrix} \mathbf{y}_0^2 \\ \eta_1^2 \end{bmatrix}$, ..., $\mathbf{y}_1^v = \begin{bmatrix} \mathbf{y}_0^v \\ \eta_1^v \end{bmatrix}$. Additionally, $\eta_1^1, \eta_1^2, \dots, \eta_1^v$ denote the outputs of the new data chunks.

We express the output weights $\bar{\beta}^{(1)}$ of t_1 as a function of $\bar{\beta}^{(0)}$, Q_1 , \bar{s}_1^p , and η_1^p , $p = 1, \dots, v$.

Similar to (9), we expand (16):

$$\begin{cases} (\bar{\mathbf{s}}_1^1)^T \mathbf{y}_1^1 = (\bar{\mathbf{s}}_0^1)^T \mathbf{y}_0^1 + (\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 \\ (\bar{\mathbf{s}}_1^2)^T \mathbf{y}_1^2 = (\bar{\mathbf{s}}_0^2)^T \mathbf{y}_0^2 + (\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 \\ \vdots \\ (\bar{\mathbf{s}}_1^v)^T \mathbf{y}_1^v = (\bar{\mathbf{s}}_0^v)^T \mathbf{y}_0^v + (\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \end{cases} \quad (17)$$

By substituting (17) into (16), we have (18). Moreover, (19) can be obtained by substituting (14) into (18). According to (19), we have the equivalent expressions shown in (20) and (21) ((18), (19) and (20) are listed at the top of next page):

$$\begin{aligned} \bar{\boldsymbol{\beta}}^{(1)} &= \frac{\bar{\boldsymbol{\beta}}^{(0)}}{\mathbf{W}} + \frac{Q_0(\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1}{\mathbf{W}} + \frac{Q_0(\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2}{\mathbf{W}} + \dots + \frac{Q_0(\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v}{\mathbf{W}} \\ &= \frac{Q_1 \bar{\boldsymbol{\beta}}^{(0)}}{Q_0} + \frac{Q_1 Q_0 (\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1}{Q_0} + \frac{Q_1 Q_0 (\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2}{Q_0} + \dots + \frac{Q_1 Q_0 (\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v}{Q_0} \\ &= \frac{Q_1 \bar{\boldsymbol{\beta}}^{(0)}}{Q_0} + Q_1 (\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 + Q_1 (\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 + \dots + Q_1 (\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \end{aligned} \quad (21)$$

where $\mathbf{W} = \mathbf{I} + Q_0 (\mathbf{s}_1^1)^T \mathbf{s}_1^1 + Q_0 (\mathbf{s}_1^2)^T \mathbf{s}_1^2 + \dots + Q_0 (\mathbf{s}_1^v)^T \mathbf{s}_1^v$. According to (13), we have

$$Q_0 (\mathbf{I} - (\mathbf{s}_1^1)^T \mathbf{s}_1^1 Q_1 - (\mathbf{s}_1^2)^T \mathbf{s}_1^2 Q_1 - \dots - (\mathbf{s}_1^v)^T \mathbf{s}_1^v Q_1) = Q_1 \quad (22)$$

$$Q_0 = \frac{Q_1}{\mathbf{I} - (\mathbf{s}_1^1)^T \mathbf{s}_1^1 Q_1 - (\mathbf{s}_1^2)^T \mathbf{s}_1^2 Q_1 - \dots - (\mathbf{s}_1^v)^T \mathbf{s}_1^v Q_1} \quad (23)$$

By substituting (23) into (21), we have the equivalent expression of the output weights shown in (24), which is listed in the next page. According to (24), we can directly update the previous global DFL model with the new output weights calculated using the data collected at time step t_1 , without retraining the existing global DFL model.

According to (11), we have

$$\begin{aligned} Q_1 &= (Q_0^{-1} + (\mathbf{s}_1^1)^T \mathbf{s}_1^1 + (\mathbf{s}_1^2)^T \mathbf{s}_1^2 + \dots + (\mathbf{s}_1^v)^T \mathbf{s}_1^v)^{-1} \\ &= \left(Q_0^{-1} + \begin{bmatrix} \mathbf{s}_1^1 \\ \mathbf{s}_1^2 \\ \vdots \\ \mathbf{s}_1^v \end{bmatrix}^T \begin{bmatrix} \mathbf{s}_1^1 \\ \mathbf{s}_1^2 \\ \vdots \\ \mathbf{s}_1^v \end{bmatrix} \right)^{-1} \end{aligned} \quad (25)$$

Based on the Woodbury formula [23], we can get

$$Q_1 = Q_0 - Q_0 \begin{bmatrix} \mathbf{s}_1^1 \\ \mathbf{s}_1^2 \\ \vdots \\ \mathbf{s}_1^v \end{bmatrix}^T \left(\mathbf{I} + \begin{bmatrix} \mathbf{s}_1^1 \\ \mathbf{s}_1^2 \\ \vdots \\ \mathbf{s}_1^v \end{bmatrix} \begin{bmatrix} \mathbf{s}_1^1 \\ \mathbf{s}_1^2 \\ \vdots \\ \mathbf{s}_1^v \end{bmatrix}^T \right)^{-1} \begin{bmatrix} \mathbf{s}_1^1 \\ \mathbf{s}_1^2 \\ \vdots \\ \mathbf{s}_1^v \end{bmatrix} Q_0 \quad (26)$$

Similar to (24) and (26), we have the general recursive formulas of updating output weights at different time steps, which are shown in (27) and (28) ((27) and (28) are listed at the top of page 7). In this way, the created MC-ELM-based DFL model can be incrementally updated using the updated output weights without retraining the existing DFL model to track environmental dynamics.

According to [23], OS-ELM updates the existing model using all the sequentially coming data equally under the assumption that those data follow the uniform distribution. Differently, MC-ELM first performs a transformation for the original objective function to separately represent the grouped

data with similar statistical characteristics using local DFL models, and then updates the existing model. In this manner, the proposed updating scheme could significantly reduce the effects of environmental dynamics on the localization performance. It should be noted that if we treat all the data equally, MC-ELM is actually equivalent to OS-ELM, thus OS-ELM is a special case of the proposed MC-ELM.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed MC-ELM-based DFL method in typical indoor environments, by comparing with selected state-of-the-art methods, including ELM, kernel ELM (K-ELM), support vector machine (SVM) [24], Kullback-Leibler (KL) divergence [25], 2D-GPR [26], random forest (RF) [27], online sequential ELM (OS-ELM) [23], FOS-ELM [28] and DU-OS-ELM [29].

A. Experimental Setup

We select a student dormitory, a laboratory, and a corridor as the testbeds. As shown in Fig. 3, both the student dormitory and laboratory belong to the typical multipath-rich indoor environments, and the multipath of the corridor also may degrade the localization performance due to its narrow layout. In the student dormitory, the number of reference points is 44, and the spacing between reference points is about 0.5m. In the laboratory, the number of reference points is 52, and the spacing between reference points is about 0.5m. In the corridor, the number of reference points is 30, and the spacing between reference points is about 0.6m. We implement a pair of WiFi transceiver in every testbed (see Fig. 3), utilizing GigaByte mini PCs with Intel 5300 NICs and external omnidirectional antennas for packet transmitting and receiving. The packet transmission rate is set as 200 packets per second, installing CSITool on the mini PCs for data collection. Moreover, we use 150 consecutive samples at each point for dataset construction.

B. Overall Comparisons in Stationary Scenarios

We first perform the performance evaluation of MC-ELM-based DFL method in stationary scenarios by comparing with ELM, K-ELM, SVM, KL, 2D-GPR and RF, and the comparison results are depicted in Fig. 4. It should be noted that the number of hidden nodes of all the local DFL models should be set as the same value for the matrix operations during the global DFL model building process. According to Fig. 4, we can find that MC-ELM-based DFL method achieves the best localization performance in the three testbeds. Specifically, both in the student dormitory and the laboratory, MC-ELM-based DFL method outperforms all the comparison methods, one of the main reasons is that MC-ELM fully considers the spatial information of the environments through dividing the whole environment into several sub-domains and representing them separately, not just treating all the data equally. Additionally, although the layout of the corridor is relatively simple, the advantages of the multidomain representation mechanism of MC-ELM are still obvious. In summary, experimental results

$$\bar{\beta}^{(1)} = Q_1 \left((\bar{s}_0^1)^T \mathbf{y}_0^1 + (\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 + (\bar{s}_0^2)^T \mathbf{y}_0^2 + (\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 + \cdots + (\bar{s}_0^v)^T \mathbf{y}_0^v + (\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \right) \quad (18)$$

$$\begin{aligned} \bar{\beta}^{(1)} &= \left(\mathbf{I} + Q_0(\mathbf{s}_1^1)^T \mathbf{s}_1^1 + Q_0(\mathbf{s}_1^2)^T \mathbf{s}_1^2 + \cdots + Q_0(\mathbf{s}_1^v)^T \mathbf{s}_1^v \right)^{-1} \\ &Q_0 \left((\bar{s}_0^1)^T \mathbf{y}_0^1 + (\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 + (\bar{s}_0^2)^T \mathbf{y}_0^2 + (\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 + \cdots + (\bar{s}_0^v)^T \mathbf{y}_0^v + (\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \right) \end{aligned} \quad (19)$$

$$\begin{aligned} &\left(\mathbf{I} + Q_0(\mathbf{s}_1^1)^T \mathbf{s}_1^1 + Q_0(\mathbf{s}_1^2)^T \mathbf{s}_1^2 + \cdots + Q_0(\mathbf{s}_1^v)^T \mathbf{s}_1^v \right) \bar{\beta}^{(1)} \\ &= Q_0 \left((\bar{s}_0^1)^T \mathbf{y}_0^1 + (\bar{s}_0^2)^T \mathbf{y}_0^2 + \cdots + (\bar{s}_0^v)^T \mathbf{y}_0^v \right) + Q_0(\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 + Q_0(\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 + \cdots + Q_0(\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \\ &= \bar{\beta}^{(0)} + Q_0(\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 + Q_0(\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 + \cdots + Q_0(\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \end{aligned} \quad (20)$$

$$\begin{aligned} \bar{\beta}^{(1)} &= \frac{\left(\mathbf{I} - Q_1(\mathbf{s}_1^1)^T \mathbf{s}_1^1 - Q_1(\mathbf{s}_1^2)^T \mathbf{s}_1^2 - \cdots - Q_1(\mathbf{s}_1^v)^T \mathbf{s}_1^v \right) Q_1 \bar{\beta}^{(0)}}{Q_1} + Q_1(\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 + Q_1(\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 + \cdots + Q_1(\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \\ &= \bar{\beta}^{(0)} - Q_1(\mathbf{s}_1^1)^T \mathbf{s}_1^1 \bar{\beta}^{(0)} - Q_1(\mathbf{s}_1^2)^T \mathbf{s}_1^2 \bar{\beta}^{(0)} - \cdots - Q_1(\mathbf{s}_1^v)^T \mathbf{s}_1^v \bar{\beta}^{(0)} + Q_1(\mathbf{s}_1^1)^T \boldsymbol{\eta}_1^1 + Q_1(\mathbf{s}_1^2)^T \boldsymbol{\eta}_1^2 + \cdots + Q_1(\mathbf{s}_1^v)^T \boldsymbol{\eta}_1^v \\ &= \bar{\beta}^{(0)} + Q_1(\mathbf{s}_1^1)^T \left(\boldsymbol{\eta}_1^1 - \mathbf{s}_1^1 \bar{\beta}^{(0)} \right) + Q_1(\mathbf{s}_1^2)^T \left(\boldsymbol{\eta}_1^2 - \mathbf{s}_1^2 \bar{\beta}^{(0)} \right) + \cdots + Q_1(\mathbf{s}_1^v)^T \left(\boldsymbol{\eta}_1^v - \mathbf{s}_1^v \bar{\beta}^{(0)} \right) \end{aligned} \quad (24)$$

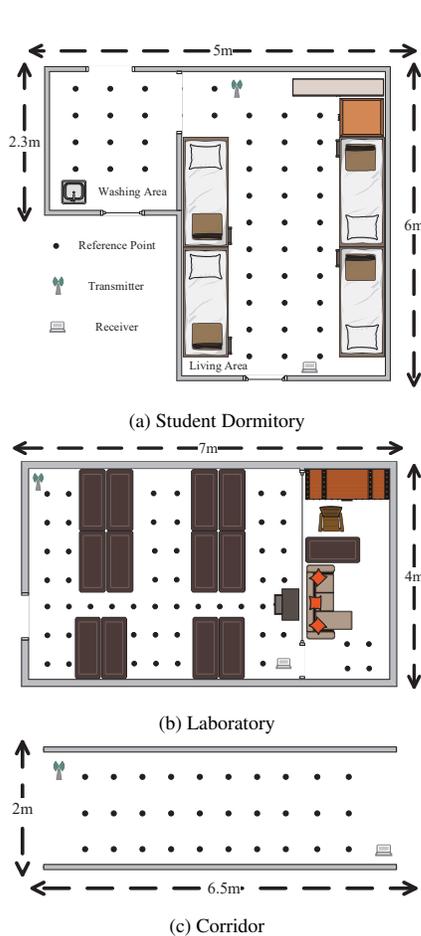


Fig. 3. Layouts of the Selected Indoor Environments

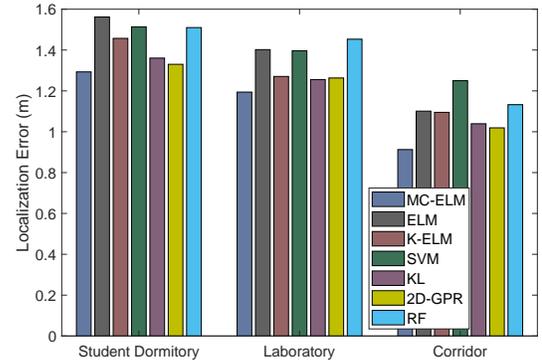


Fig. 4. Comparison Results of Stationary Scenarios

in stationary scenarios demonstrate the robustness of MC-ELM-based DFL method in cluttered environments.

In order to further evaluate the validity of the domain decomposition of MC-ELM in performance improvement, we divide the student dormitory and laboratory into several sub-domains depending on their layouts, and then build a corresponding number of individual DFL models for all the sub-domains and a general DFL model for the two testbeds using ELM. Specifically, the student dormitory is divided into two sub-domains, including the washing area and the living area, and the laboratory is divided into three sub-domains, respectively. Table 1 lists the detailed comparison results, in which individual DFL models achieve better localization performance than the corresponding general DFL models. The main reason is that individual DFL models are built using the data with uniform distributions, but the general DFL models utilize all the data collected from the whole

$$Q_{k+1} = Q_k - Q_k \begin{bmatrix} \mathbf{s}_{k+1}^1 \\ \mathbf{s}_{k+1}^2 \\ \vdots \\ \mathbf{s}_{k+1}^v \end{bmatrix}^T \left(\mathbf{I} + \begin{bmatrix} \mathbf{s}_{k+1}^1 \\ \mathbf{s}_{k+1}^2 \\ \vdots \\ \mathbf{s}_{k+1}^v \end{bmatrix} Q_k \begin{bmatrix} \mathbf{s}_{k+1}^1 \\ \mathbf{s}_{k+1}^2 \\ \vdots \\ \mathbf{s}_{k+1}^v \end{bmatrix}^T \right)^{-1} \begin{bmatrix} \mathbf{s}_{k+1}^1 \\ \mathbf{s}_{k+1}^2 \\ \vdots \\ \mathbf{s}_{k+1}^v \end{bmatrix} Q_k \quad (27)$$

$$\bar{\beta}^{(k+1)} = \bar{\beta}^{(k)} + Q_{k+1} (\mathbf{s}_{k+1}^1)^T \left(\eta_{k+1}^1 - \mathbf{s}_{k+1}^1 \bar{\beta}^{(k)} \right) + Q_{k+1} (\mathbf{s}_{k+1}^2)^T \left(\eta_{k+1}^2 - \mathbf{s}_{k+1}^2 \bar{\beta}^{(k)} \right) + \dots + Q_{k+1} (\mathbf{s}_{k+1}^v)^T \left(\eta_{k+1}^v - \mathbf{s}_{k+1}^v \bar{\beta}^{(k)} \right) \quad (28)$$

TABLE I
COMPARISON RESULTS OF INDIVIDUAL DFL MODELS AND GENERAL DFL MODEL

DFL Model	Student Dormitory	Laboratory
Individual Model 1	1.4127	1.2930
Individual Model 2	1.3009	1.2105
Individual Model 3	–	1.3362
General Model	1.5531	1.4096

TABLE II
COMPARISON RESULTS OF THE NONSTATIONARY SCENARIOS WITH CHUNK-BY-CHUNK COMING DATA

Method	Chunk Size	Localization Error (m)
MC-ELM	5-by-5	1.2109
	5-by-15	1.2565
	[10, 30]	1.2011

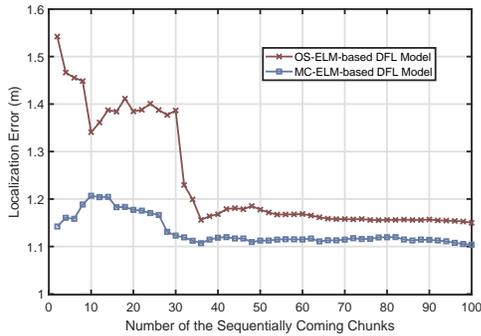


Fig. 5. Convergence Comparison between MC-ELM-based DFL Model and OS-ELM-based DFL Model

environment, which may follow different distributions, leading to performance loss.

C. Overall Comparisons in Nonstationary Scenarios

In this section, we evaluate the localization performance of MC-ELM-based DFL method in nonstationary scenarios by comparing with OS-ELM, FOS-ELM and DU-OS-ELM. We first show the convergence of MC-ELM-based DFL model, and the detailed comparison results are depicted in Fig. 5. According to Fig. 5, we can find that the localization error of MC-ELM decreases with the increase of the sequentially coming data, it indicates that MC-ELM could adapt to the new environment by updating the existing DFL model using the newly collected data. Moreover, the convergence speed of MC-ELM is much faster than OS-ELM, demonstrating its great capability to track environmental dynamics.

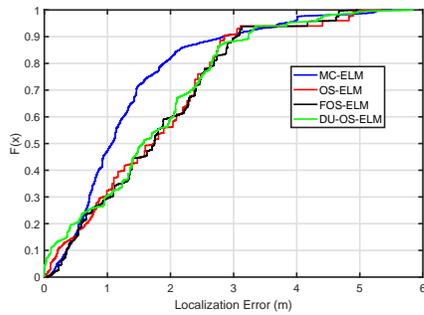
Fig. 6 illustrates the cumulative distribution function (CDF) of the three testbeds with one-by-one coming data collected from other time steps, in which MC-ELM-based DFL method achieves the best localization performance in all the situations. Additionally, Table II lists the comparison results of various

chunk-by-chunk learning modes. Specifically, the fixed data chunk size of 5 and 15, as well as a randomly varying data chunk size between 10 and 30 are implemented. According to Table II, the chunk sizes of the sequentially coming data have no obvious effects on localization performance of MC-ELM. It should be noted that although the data collected from other time steps may follow different distributions, the proposed MC-ELM still could keep the localization performance through incrementally updating the existing DFL model with the newly coming data, indicating its excellent dynamic tracking capability in nonstationary scenarios.

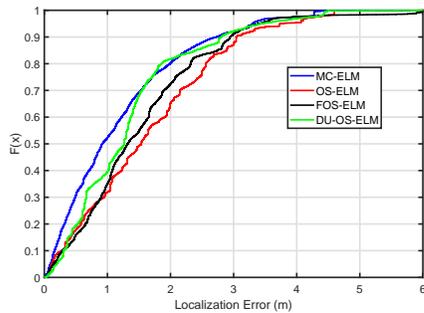
Remark 1 DFL in nonstationary scenarios means that the DFL model is initially built with the data collected at one specific time step, and then the newly collected data are coming sequentially chunk-by-chunk. The sequentially coming data may follow different distributions comparing with the initially collected data, due to the seasonality or periodicity effects (such as opening/closing window/door), function faults of the sensors embedded in WiFi devices, and thermal drifts of WiFi devices, etc.

D. Effects of Packet Transmission Rates

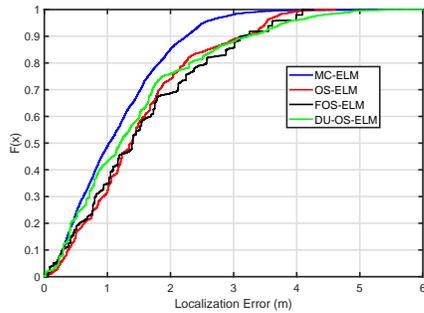
In this section, we explore the effects of the packet transmission rate on the localization performance. Fig. 7 illustrates the comparison results of MC-ELM-based DFL method with various packet transmission rates. According to Fig. 7, the localization performance is relatively worse with the packet transmission rates 50 and 100, but it becomes better gradually with the increasing packet transmission rate. It indicates that low packet transmission rate may not capture the target's location information well, and higher packet transmission rate could guarantee the localization performance, but easily lead to time-consuming data processing and model building. In practice, we should balance the localization performance and time consumption. Moreover, Fig. 8 shows the packet loss rate when packet transmission rate is 300. Accordingly, we can find that there seems no serious packet loss using our WiFi devices.



(a) CDF of the Localization Performance in the Student Dormitory



(b) CDF of the Localization Performance in the Laboratory



(c) CDF of the Localization Performance in the Corridor

Fig. 6. Comparison Results of Nonstationary Scenarios with One-by-One Coming Data

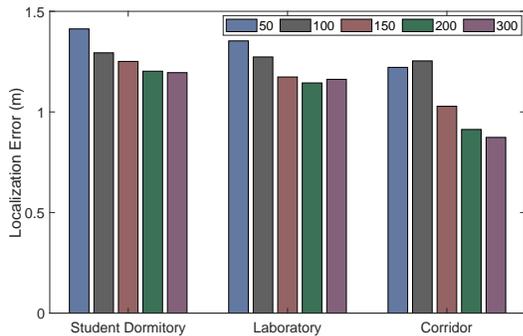


Fig. 7. Localization Performance with Various Packet Transmission Rates

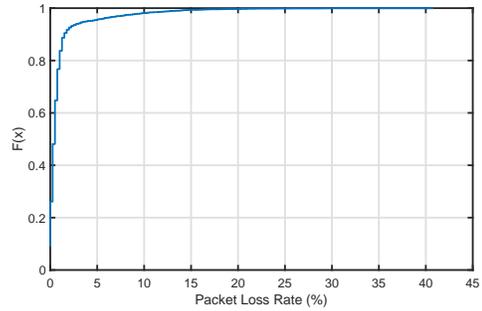


Fig. 8. CDF of the Packet Loss Rate

V. DISCUSSIONS

In this section, we discuss several issues of this work, including the essence of DFL in nonstationary environments, some limitations and future directions.

A. Understanding DFL in Nonstationary Scenarios

Most existing DFL methods belong to batch learning methods, which ignore the environmental dynamics, such as layout of the environment, deployment of WiFi devices, thermal drifts of WiFi devices, and variations of humidity, etc. However, in practical scenarios, the environment is nonstationary and time-varying. The batch learning methods are time-consuming and labor-intensive for model updating in dealing with environmental dynamics. Differently, MC-ELM-based DFL method can directly update the existing DFL model with the received new data, and does not require retraining. Moreover, traditional incremental learning methods only can handle the one-by-one coming data with specific types of hidden nodes, such as growing and pruning RBF network (GAP-RBF) and stochastic gradient descent back-propagation (SGBP) [30]. MC-ELM can handle both one-by-one and chunk-by-chunk coming data with fixed or varying size, making MC-ELM-based DFL model more flexible and practical in model updating.

B. Limitations and Future Directions

In nonstationary scenarios, sequentially coming data usually carry more useful information about the current environment, and the historical data collected from previous time steps even may contain redundant information, because the data usually have timeliness. However, MC-ELM treats both historical data and new data equally. Furthermore, the correction term of online learning algorithm is losing its capability with accumulated data, but MC-ELM cannot automatically measure whether the sequentially coming data are necessary for DFL model updating. Thus, it is worthwhile to introduce data filtering mechanism before model updating, such as the approximate linear dependency (ALD), as well as forgetting mechanism for the historical data, making the DFL model more adaptive in nonstationary scenarios.

VI. CONCLUSIONS

In this paper, we propose a modified MC-ELM-based DFL framework to enhance the localization performance in nonstationary environments. Specifically, the whole environment

is first divided into several sub-domains depending on the distributions of the collected data. After that, these sub-domains are represented using a corresponding number of local DFL models separately. Finally, a global DFL model is built through seamlessly integrating all the local DFL models in a global optimization manner. MC-ELM-based DFL framework also can incrementally update the existing DFL model with sequentially coming data without retraining. In this way, MC-ELM-based DFL framework is not only robust by fully considering the spatiotemporal information of the environment, but also lightweight in dealing with environmental dynamics. The proposed MC-ELM-based DFL framework outperforms selected state-of-the-art methods in several indoor environments, indicating its robustness, great generalization performance, and good dynamic tracking capability.

VII. APPENDIX

In order to solve the objective function demonstrated in (3), we first transform (3) as the following equivalent expressions:

$$\begin{aligned} \min : & \frac{1}{2} \sum_{p=1}^v \left\| \beta_p^{(0)} \right\|^2 + \frac{1}{2} \lambda \sum_{p=1}^v \sum_{q=1}^{\tilde{N}_p^0} (\varsigma_{pq}^0)^2 + \\ & \frac{1}{2} \mu \frac{v^2}{(v-1)^2} \sum_{p=1}^v \left\| \beta_p^{(0)} - \frac{1}{v} \sum_{l=1}^v \beta_l^{(0)} \right\|^2 \\ \text{s.t.}, & s_0(\mathbf{x}_{pq}^0) \beta_p^{(0)} = \mathbf{y}_{pq}^0 - \varsigma_{pq}^0, p = 1, \dots, v, q = 1, \dots, \tilde{N}_p^0 \end{aligned} \quad (29)$$

Then, we have

$$\begin{aligned} \min : & \frac{1}{2} \sum_{p=1}^v \left\| \beta_p^{(0)} \right\|^2 + \frac{1}{2} \lambda \sum_{p=1}^v \sum_{q=1}^{\tilde{N}_p^0} (\varsigma_{pq}^0)^2 + \frac{1}{2} \mu \sum_{p=1}^v \left\| \beta_p^{(0)} - \hat{\beta}^{(0)} \right\|^2 \\ \text{s.t.}, & s_0(\mathbf{x}_{pq}^0) \beta_p^{(0)} = \mathbf{y}_{pq}^0 - \varsigma_{pq}^0, p = 1, \dots, v, q = 1, \dots, \tilde{N}_p^0 \end{aligned} \quad (30)$$

where $\hat{\beta}^{(0)} = \frac{1}{v} \sum_{l=1}^v \beta_l^{(0)}$ represents the global output weights.

Let $\tilde{\beta}_p^{(0)} = \beta_p^{(0)} - \hat{\beta}^{(0)}$, the above objective function is

$$\begin{aligned} \min : & \frac{1}{2} \sum_{p=1}^v \left\| \tilde{\beta}_p^{(0)} + \hat{\beta}^{(0)} \right\|^2 + \frac{1}{2} \lambda \sum_{p=1}^v \sum_{q=1}^{\tilde{N}_p^0} (\varsigma_{pq}^0)^2 + \frac{1}{2} \mu \sum_{p=1}^v \left\| \tilde{\beta}_p^{(0)} \right\|^2 \\ \text{s.t.}, & s_0(\mathbf{x}_{pq}^0) (\tilde{\beta}_p^{(0)} + \hat{\beta}^{(0)}) = \mathbf{y}_{pq}^0 - \varsigma_{pq}^0, p = 1, \dots, v, q = 1, \dots, \tilde{N}_p^0 \end{aligned} \quad (31)$$

Let $\bar{\beta}^{(0)} = \left[\left(\tilde{\beta}_1^{(0)} + \hat{\beta}^{(0)} \right), \left(\tilde{\beta}_2^{(0)} + \hat{\beta}^{(0)} \right), \dots, \left(\tilde{\beta}_v^{(0)} + \hat{\beta}^{(0)} \right), \sqrt{\mu} \tilde{\beta}_1^{(0)}, \sqrt{\mu} \tilde{\beta}_2^{(0)}, \dots, \sqrt{\mu} \tilde{\beta}_v^{(0)} \right]$, and we can achieve the final equivalent objective function of MC-ELM, which is shown in (4).

REFERENCES

- [1] Y. Ma, G. Zhou, and S. Wang, "Wi-Fi sensing with channel state information: A survey," *ACM Comput. Surv. Tut.*, vol. 1, no. 1, pp. 1-35, 2019.
- [2] Y. Ma, B. Wang, X. Gao, and W. Ning, "The gray analysis and machine learning for device-free multitarget localization in passive UHF RFID environments," *IEEE Trans. Ind. Inform.*, vol. 16, no. 2, pp. 802-813, 2020.
- [3] X. Guo, R. Nkrow, N. Ansari, L. Lin, and L. Wang, "Robust WiFi localization by fusing derivative fingerprints of RSS and multiple classifiers," *IEEE Trans. Ind. Inform.*, vol. 16, no. 5, pp. 3177-3186, 2020.
- [4] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 1-31, 2013.
- [5] M. Nabati, H. Navidan, R. Shahbazian, and S.A. Ghorashi, "Using synthetic data to enhance the accuracy of fingerprint-based localization: A deep learning approach," *IEEE Sens. Lett.*, vol. 4, no. 4, pp. 1-4, 2020.
- [6] J. Zhang, W. Xiao, and Y. Li, "Integrated multiple kernel learning for device-free localization in cluttered environments using spatiotemporal information," *IEEE Internet of Things J.*, vol. 8, no. 6, pp. 4749-4761, 2021.
- [7] T. Xin, B. Guo, Z. Wang, P. Wang, J.C.K. Lam, V. Li, and Z. Yu, "FreeSense: A robust approach for indoor human detection using Wi-Fi signals," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 3, 2018, pp. 1-23.
- [8] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: A survey," *IEEE Commun. Surv. Tut.*, vol. 22, no. 3, pp. 1629-1645, 2020.
- [9] J. Wang, Q. Gao, Y. Yu, X. Zhang, and X. Feng, "Time and energy efficient TOF-based device-free wireless localization," *IEEE Trans. Ind. Inform.*, vol. 12, no. 1, pp. 158-168, 2016.
- [10] Y. Li, D. Wu, J. Zhang, X. Xu, Y. Xie, T. Gu, and D. Zhang, "DiverSense: Maximizing Wi-Fi sensing range leveraging signal diversity," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 6, no. 2, 2022, pp. 1-28.
- [11] J. Zhang, Y. Li, X. Xiong, D. Dou, C. Miao, and D. Zhang, "HandGest: Hierarchical sensing for robust in-the-air handwriting recognition with commodity WiFi devices," *IEEE Internet of Things J.*, vol. 9, no. 19, 2022, pp. 19529-19544.
- [12] Y. Yang, J. Cao, X. Liu, and X. Liu, "Door-monitor: Counting in-and-out visitors with COTS WiFi devices," *IEEE Internet of Things J.*, vol. 7, no. 3, pp. 1704-1717, 2019.
- [13] Y. Wang, K. Wu, and L.M. Ni, "WiFall: Device-free fall detection by wireless networks," *IEEE Trans. Mob. Comput.*, vol. 16, no. 2, pp. 581-594, 2016.
- [14] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *ACM SIGCOMM Computer Commun. Rev.*, vol. 41, no. 1, pp. 53, 2011.
- [15] J. Wang, Q. Gao, Y. Yu, P. Cheng, L. Wu, and H. Wang, "Robust device-free wireless localization based on differential RSS measurements," *IEEE Trans. Ind. Electron.*, vol. 60, no. 12, pp. 5943-5952, 2013.
- [16] K. Ohara, T. Maekawa, Y. Kishino, Y. Shirai, and F. Naya, "Transferring positioning model for device-free passive indoor localization," in *Proc. 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, Osaka, Japan, pp. 885-896, 2015.
- [17] L. Chang, X. Chen, Y. Yang, D. Fang, J. Wang, T. Xing, and Z. Tang, "FitLoc: Fine-grained and low-cost device-free localization for multiple targets over various areas," *IEEE Trans. Netw.*, vol. 25, no. 4, pp. 1994-2007, 2017.
- [18] Q. Wang, X. Yin, J. Tan, T. Xing, J. Niu, and Z. Tang, "DTransfer: Extremely low cost localization irrelevant to targets and regions for activity recognition," *Pers. Ubiquit. Comput.*, vol. 23, no. 1, pp. 3-16, 2019.
- [19] 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 J.*, vol. 7, no. 12, pp. 11851-11862, 2020.
- [20] J. Wang, X. Chen, D. Fang, C.Q. Wu, Z. Yang, and T. Xing, "Transferring compressive-sensing-based device-free localization across target diversity," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2397-2409, 2015.
- [21] G. Ditzler, M. Roveri, C. Alippi, and R. Polikar, "Learning in nonstationary environments: A survey," *IEEE Comput. Intell. Mag.*, vol. 10, no. 4, pp. 12-25, 2015.
- [22] J. Zhang, Y. Li, W. Xiao, and Z. Zhang, "Non-iterative and fast deep learning: Multilayer extreme learning machines," *J. Franklin Inst.*, vol. 357, no. 13, pp. 8925-8955, 2020.
- [23] N.Y. Liang, G.B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Trans. Neural Netw.*, vol. 17, no. 6, pp. 1411-1423, 2006.
- [24] R. Zhou, X. Lu, P. Zhao, and J. Chen, "Device-free presence detection and localization with SVM and CSI fingerprinting," *IEEE Sens. J.*, vol. 17, no. 23, pp. 7990-7999, 2017.
- [25] R. Gao, J. Zhang, W. Xiao, and Y. Li, "Kullback-Leibler divergence based probabilistic approach for device-free localization using channel state information," *Sensors*, vol. 19, no. 11, pp. 4783-4806, 2019.
- [26] M. Nabati, S.A. Ghorashi, and R. Shahbazian, "Joint coordinate optimization in fingerprint-based indoor positioning," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1192-1195, 2021.

- [27] D.J. Suroso, R. Rupaksi, A.B. Krisnawan, and N.A. Siddiq, "Random forest-based fingerprinting technique for device-free indoor localization system," *Indones. J. Comput. Eng. Design*, vol. 3, no. 2, pp. 79-96, 2021.
- [28] J. Zhao, Z. Wang and D.S. Park, "Online sequential extreme learning machine with forgetting mechanism," *Neurocomputing*, vol. 87, pp. 79-89, 2012.
- [29] Y. Li, S. Zhang, Y. Yin, W. Xiao, and J. Zhang, "A novel online sequential extreme learning machine for gas utilization ratio prediction in blast furnaces," *Sensors*, vol. 17, no. 8, pp. 1847-1870, 2017.
- [30] Y. LeCun, L. Bottou, G.B. Orr, and K.R. Muller, "Efficient backprop," *Lecture Notes Comput. Sci.*, vol. 1524, pp. 9-50, 1998.