DHCLoc: A Device Heterogeneity Tolerant and Channel Adaptive Passive WiFi Localization Method Based on DNN

Lifei Hao, Baoqi Huang, *Member, IEEE,* Bing Jia and Guoqiang Mao, *Fellow, IEEE*

***Abstract*—Passive WiFi localization refers to determining the location of WiFi enabled mobile devices by deploying dedicated WiFi access points to sniff WiFi packets transmitted by these mobile devices and measure the corresponding Received Signal Strengths (RSSs) for use in localization. However, most existing studies fail to consider the effect of multiple channels where WiFi packets are transmitted and sniffed. The problem is further exacerbated by device heterogeneity occurring cross various mobile devices. In this paper, we present a unified Deep Neural Network (DNN)-based solution, termed DHCLoc, to address these two challenges. To be specific, a Cramer-Rao Lower Bound (CRLB)-based analysis reveals that utilizing multi-channel information will benefit localization, motivating us to include channel information into DHCLoc. Moreover, a novel Maximum Likelihood Estimation (MLE)-based localization framework is introduced by incorporating a new variable to characterize the RSS measurement offsets caused by device heterogeneity, inspiring us to apply adversarial training to adopt such offsets against device heterogeneity. Extensive experiments using two real-world datasets are conducted, and show that, in comparison with several existing methods, DHCLoc can improve the localization accuracy by at least** 25.2% **and** 25.8% **respectively.**

***Index Terms*—WiFi Localization, RSS, channel, device heterogeneity, deep learning.**

I. Introduction

NOWADAYS, there is an increasing demand on Location Based Services (LBSs) [2], and various solutions, such as bluetooth [3], WiFi fingerprinting [4] and Channel State Information (CSI) [5] etc., have been developed. Therein, WiFi fingerprinting is one of the most popular solutions due to its wide availability and low costs.

The basic process of WiFi fingerprinting involves two stages [6], i.e., offline training and online matching. In the offline stage, Received Signal Strength (RSS) measurements from different WiFi access points (APs) are collected at different reference points (RPs), producing an RSS fingerprint database (FD). In the online stage, the RSS measurement vector obtained at an unknown position is matched across the FD, such that one optimal RP is selected. As to fingerprint matching algorithms, various methods like k-Nearest Neighbors (KNN) [7], weighted KNN (WKNN) [8] and support vector machines (SVM) [9] have been proposed. However, these methods suffer from their limited ability of learning complex features from training samples. Recent works focused on using deep learning models [10]–[14] due to their powerful learning ability [15].

There are two WiFi positioning modes, i.e., active mode and passive mode [16]. In the active mode, mobile devices collect RSS measurements from nearby APs through WiFi active scans, which is suitable for the cases where smartphones proactively trigger localization. Conversely, in the passive mode, the WiFi sniffing technique is employed to collect WiFi packets sent by nearby mobile devices, such that the corresponding RSS measurements can be utilized for localization [4], which is suitable for the cases where localization is triggered and used by server-side systems. Specifically, multiple WiFi modules can be installed in one WiFi AP to simultaneously monitor more than one WiFi channel, with the result that the influences of packet collisions and channel interferences can be mitigated. But signals transmitted with the same power in different channels (i.e. with different frequencies) will incur different path losses, thus producing different RSS measurements at a fixed receiver and consequently degrading localization performance, which is even worse when both $2.4G$ and $5G$ bands are involved [17]. However, existing studies fail to address this issue.

Furthermore, mobile devices used for constructing an FD are often different from those in the online stage, which is dubbed the problem of device heterogeneity. Consequently, RSS measurements obtained by heterogeneous devices at the same location usually have different mean values [18], and might be translated into different physical locations by WiFi fingerprinting localization methods [19]–[21]. Previous researches focused on developing new localization algorithms or constructing new types of FD. However, with the emergence of various brands and models of mobile devices, this issue would become more pronounced, and it is imperative to develop robust localization methods.

To tackle the above challenges, we present a Deep Neural Network (DNN)-based passive WiFi fingerprinting localization algorithm, termed DHCLoc, which takes into account channel information associated with each RSS measurement and the offsets in RSS measurements from heterogeneous devices. To be specific, we investigate the effect of channel frequencies on RSS measurements from both theoretical and experimental

L. Hao, B. Huang and B. Jia are with the College of Computer Science, Inner Mongolia University, and the Inner Mongolia A.R. Key Laboratory of Wireless Networking and Mobile Computing, Hohhot 010021, China.

G. Mao is with the Research Institute of Smart Transportation, Xidian University, Xi'an 710071, China.

Corresponding Author: Baoqi Huang, E-mail: cshbq@imu.edu.cn

Part of this work has been published in a conference proceeding [1].

Manuscript received Xxx xx, 20xx; revised Xxx xx, 20xx.

perspectives, and confirm the superiority of utilizing multi-channel RSS measurements for localization. Moreover, we propose an optimal offset estimation (OOE) method based on Maximum Likelihood Estimation (MLE) to infer different offsets incurred by various mobile devices, and further apply the adversarial training to combat device heterogeneity. On these grounds, a DNN-based localization model is designed. Extensive experiments conducted in two scenarios demonstrate the effectiveness of multi-channel RSS utilization and OOE, and indicate that DHCLoc can significantly outperform several baseline algorithms by at least 25.2% and 25.8% with the two datasets, respectively. The main contributions of this work can be summarized as follows:

- We propose to exploit the channel information to improve the performance of passive WiFi fingerprinting localization, which was omitted by previous studies. In addition, a theoretical analysis is provided.
- Unlike existing methods explicitly defining new fingerprints, we present an innovative method to combat device heterogeneity, which can be easily embedded into existing localization algorithms.
- Field experiments are conducted for validation, and two datasets are shared on GitHub[1].

The rest of this paper is organized as follows. Next section introduces related works. In Sections III, IV and V, the theoretical analysis of localization with multi-channel RSS measurements, OOE and DNN-based localization model are described, respectively. Our work is validated in Section VI through extensive experiments. The final section summarizes the paper and sheds lights on future works.

II. Related Work

In this section, we briefly review the literatures related to our study.

A. WiFi Fingerprinting Localization

Since RADAR [22] opened up the field of WiFi fingerprinting-based localization, various techniques have been reported [7], [8], [23]. Recently, some works based on machine learning approaches, including SVM [9], factor graphs [24], kernel estimation [25], extreme learning machine (ELM) [20], [26] and etc., have attracted attentions due to their fast learning and easy implementation. The state-of-the-art researches mainly focus on the following two directions.

1) Improving the Expression of Fingerprints (or FDs): To enhance the WiFi fingerprinting approaches, various studies seek robust RSS fingerprints [23]. GIFT [27] instead introduced binary differential RSS measurements as fingerprints; ViVi [28] and ViViPlus [29] explored the gradient of selected multiple neighboring fingerprints to deal with spatial ambiguities. Another line of works focused on improving the efficiency of FD. For instance, in [4], an adaptive weighted fusion algorithm based on fingerprint recognition and trilateration was proposed; Huang et al. proposed an online radio map update scheme [30] by adopting extended Gaussian Process Regression (GPR) to alleviate the model inaccuracy due to noisy location labels. Though these works have attained successes to a certain extent, researchers are still interested in developing advanced localization algorithms in order to achieve superior performance.

2) Deep Learning based Localization: Some recent works adopted deep learning models to enhance fingerprint matching. For instance, WiDeep [10] integrates a stacked auto-encoders model with a probabilistic framework to handle noises and capture complex relationships among WiFi signals; Belay et al. constructed a database including RSS measurements and corresponding AP identifiers, applied linear discriminant analysis to extract features, and utilized a DNN model to predict locations [11]; Xiong et al. devised an artificial multi-layer neural network using channel impulse responses as fingerprints [12]. In addition, as a variant of DNN, the CNN-based WiFi fingerprinting localization was presented to pursue better performance [13], [14]. These deep learning based methods, though can effectively improve localization accuracy, still suffer from high training costs and model transferring.

B. Device Heterogeneity Mitigation

To handle the device heterogeneity issue, different schemes have been reported [19], [20], [31]–[34].

One effective but time-consuming solution is to manually adjust RSS measurements for distinct testing devices via a linear transformation method [31], [32]. Its main drawback lies in requiring the knowledge of all heterogeneous devices in advance. Some calibration-free methods were proposed in [33], [34] to avert the tedious manual calibration efforts for each testing device. Nevertheless, these methods rely on time-consuming online processing.

Another solution is to define alternative location fingerprints. Signal strength difference (SSD) defines the difference between RSS measurements as a location fingerprint [19], but suffers from the effect of shadowing variation and information loss caused by reducing the dimensionality of fingerprints. In [20], standardized fingerprints were introduced based on Procrustes Analysis (PA), but since only the information involved in one fingerprint is utilized for standardizing itself and the characteristics cross the whole FD are ignored, the resulting performance is quite limited.

However, different from these existing approaches, we do not assume to know all mobile devices in advance and aim to explore as much information in a FD as possible to combat device heterogeneity.

III. Multi-channel Information Utilization

In this section, we investigate the influence of different channel frequencies on RSS measurements, and then conduct a comprehensive analysis with multi-channel RSS measurements based on Cramer-Rao Lower Bound (CRLB).

A. Influence of Channel Frequencies on RSS Measurements

In the literature, various theoretical models have been established to characterize the propagation of wireless signals.

[1]The datasets have been submitted to GitHub: https://github.com/vtcm800/MultiChannelHeterogeneityDNNExperiments

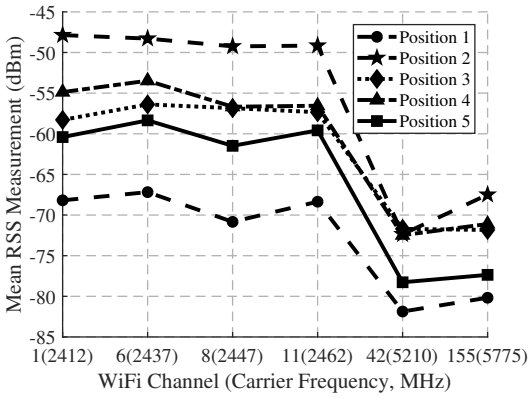


Fig. 1. Mean RSS measurements from 6 different WiFi channels in five random positions of the OFFICE dataset.

Specifically, both the Friis free space model and the more realistic log-normal path loss model indicates that RSS is relevant to carrier frequency [35], [36], namely that the higher is carrier frequency, the larger is path loss, and thus the smaller is the resulting RSS.

The above theoretical analysis is also confirmed by our experimental studies. In Fig. 1, the mean RSS measurements between the same pair of transmitter and receiver in different WiFi channels are plotted. It can be seen that, the mean RSS measurement generally decreases with the frequency increasing, especially when the channel rises from $11(2451MHz)$ to $42(5210MHz)$ since there exists a huge change in the frequency.

In summary, it is noticeable that the carrier frequency has significant influences on RSS measurements, inspiring us to pay attention to such valuable information.

B. Theoretical Analysis of Localization Errors with Multi-Channel Measurements

Suppose that there are c channels, and in each channel, m APs obtain m RSS measurements from one mobile device. According to the usage of multi-channel RSS measurements, we can formulate two CRLBs as follows.

In the first case, the fingerprint obtained from a mobile device at an arbitrary position $\mathbf{l} = [l_1, l_2]^T$ is formulated as $\mathbf{r} = [r_{11}, r_{12}, ..., r_{ij}, ..., r_{cm}]^T$. By adopting the similar assumption as in [17], [37], $\mathbf{r}$ is assumed to be independent and identically distributed (i.i.d.), i.e.

$$\mathbf{r} \sim N(\mathbf{m}(\mathbf{l}), \sigma^2 \mathbf{I}_{cm}), \tag{1}$$

where $\mathbf{m}(\mathbf{l}) = [m_{11}(\mathbf{l}), m_{12}(\mathbf{l}), ..., m_{cm}(\mathbf{l})]^T$ is the vector function containing mean functions of RSS measurements in relation to different APs and channels, and $\mathbf{I}_{cm}$ is a $c \times m$ identity matrix.

When using MLE to infer this unknown position $\mathbf{l}$, the corresponding log likelihood function is defined as

$$L(\mathbf{r}; \mathbf{l}) = \log p(\mathbf{r}|\mathbf{l}), \tag{2}$$

and thus, the Fisher information matrix (FIM) in the multi-channel case, denoted $\mathbf{F}_{mc}(\mathbf{l})$, is

$$\mathbf{F}_{mc}(\mathbf{l}) = -E\left(\frac{\partial^2 L(\mathbf{r}; \mathbf{l})}{\partial \mathbf{l} \partial \mathbf{l}^T}\right) = \frac{1}{\sigma^2}\mathbf{g}^T\mathbf{g}, \tag{3}$$

where $\mathbf{g} = [\mathbf{g}_{11}^T, \mathbf{g}_{12}^T, ..., \mathbf{g}_{cm}^T]^T$, and $\mathbf{g}_{ij} = \frac{\partial m_{ij}}{\partial \mathbf{l}} = [g_1^{ij}, g_2^{ij}]$ denotes the gradient representing the changing degrees of the mean RSS in the ith channel from the jth AP.

Thus, the CRLB is calculated as

$$\begin{aligned}\mathbf{F}_{mc}^{-1}(\mathbf{l}) &= \sigma^2 \left[\sum_{i=1}^{c}\sum_{j=1}^{m}\begin{pmatrix}(g_1^{ij})^2 & g_1^{ij}g_2^{ij}\\ g_1^{ij}g_2^{ij} & (g_2^{ij})^2\end{pmatrix}\right]^{-1} \\ &= \frac{\sigma^2 \sum_i \sum_j \begin{pmatrix}(g_1^{ij})^2 & g_1^{ij}g_2^{ij}\\ g_1^{ij}g_2^{ij} & (g_2^{ij})^2\end{pmatrix}}{\sum_i \sum_j (g_1^{ij})^2 \sum_i \sum_j (g_2^{ij})^2 - (\sum_i \sum_j g_1^{ij} g_2^{ij})^2}.\end{aligned} \tag{4}$$

Finally, the mean squared error (MSE), i.e. the trace of CRLB, is derived as follows

$$\begin{aligned}\mathrm{Tr}(\mathbf{F}_{\mathrm{mc}}^{-1}(\mathbf{l})) &= \frac{\sigma^2 \sum_i \sum_j ((g_1^{ij})^2 + (g_2^{ij})^2)}{\sum_i \sum_j (g_1^{ij})^2 \sum_i \sum_j (g_2^{ij})^2 - (\sum_i \sum_j g_1^{ij} g_2^{ij})^2} \\ &= \frac{2\sigma^2 \sum_i \sum_j \|\mathbf{g}_{ij}\|^2}{\sum_i \sum_j \sum_p \sum_q (\|\mathbf{g}_{ij}\|\|\mathbf{g}_{pq}\| sin\theta_{ij}^{pq})^2} \\ &= \frac{2\sigma^2}{\sum_i \sum_j \left(\frac{\|\mathbf{g}_{ij}\|^2}{\sum_k \sum_r \|\mathbf{g}_{kr}\|^2} \sum_i \sum_j (h_{ij}^{pq})^2\right)} \\ &= \frac{2\sigma^2}{cm \sum_i \sum_j \left(\frac{\|\mathbf{g}_{ij}\|^2}{\sum_k \sum_r \|\mathbf{g}_{kr}\|^2} (\overline{h}_{ij}^{mc})^2\right)} = \frac{2\sigma^2}{cm(\overline{h}_w^{mc})^2},\end{aligned} \tag{5}$$

where θ_{ij}^{pq} is the angle subtended by $\mathbf{g}_{ij}$ and $\mathbf{g}_{pq}$,

$$h_{ij}^{pq} = \|\mathbf{g}_{pq}\| sin\theta_{ij}^{pq}, \tag{6}$$

$$\overline{h}_{ij}^{mc} = \sqrt{\frac{1}{cm}\sum_{p=1}^{c}\sum_{q=1}^{m}(h_{ij}^{pq})^2}, \tag{7}$$

$$\overline{h}_w^{mc} = \sqrt{\sum_{i=1}^{c}\sum_{j=1}^{m}\left(\frac{\|\mathbf{g}_{ij}\|^2}{\sum_k \sum_r \|\mathbf{g}_{kr}\|^2}(\overline{h}_{ij}^{mc})^2\right)}. \tag{8}$$

It is clear that $(\overline{h}_w^{mc})^2$ in (8) and $(\overline{h}_{ij}^{mc})^2$ in (7) respectively denote a weighted average and an average, reflecting that $(\overline{h}_w^{mc})^2$ generally does not scale with c. Therefore, it can be concluded that the MSE is inversely proportional to c and m. In particular, when $c = 1$, it is equivalent to the original case in [17], i.e. localization with single channel measurements.

In the second case, the multi-channel information is ignored by averaging the multi-channel RSS measurements from one AP to form one element of a fingerprint, as the traditional method does. The average RSS measurement, denoted $\mathbf{r}^{mix}$ will satisfy the following Gaussian distribution

$$\mathbf{r}^{mix} \sim N(\mathbf{m}^{mix}(\mathbf{l}), \frac{\sigma^2}{c}\mathbf{I}_m), \tag{9}$$

where $\mathbf{m}^{mix}(\mathbf{l}) = [\sum_{i=1}^{c} m_{i1}(\mathbf{l}), \sum_{i=1}^{c} m_{i2}(\mathbf{l}), ..., \sum_{i=1}^{c} m_{im}(\mathbf{l})]$.

By letting θ_j^q be the angle subtended by $\mathbf{g}_j$ and $\mathbf{g}_q$ with

$$\mathbf{g}_j = \frac{\sum_{i=1}^{c} \mathbf{g}_{ij}}{c}, \tag{10}$$

the MSE can be similarly obtained as

$$\mathrm{Tr}(\mathbf{F}_{\mathrm{mix}}^{-1}(\mathbf{l})) = \frac{2\sigma^2}{\mathrm{cm}(\overline{\mathrm{h}}_{\mathrm{w}}^{\mathrm{mix}})^2}, \tag{11}$$

with

$$\overline{h}_w^{mix} = \sqrt{\sum_{j=1}^{m} \left(\frac{\|\mathbf{g}_j\|^2}{\sum_r \|\mathbf{g}_r\|^2} (\overline{h}_j^{mix})^2 \right)}, \tag{12}$$

$$\overline{h}_j^{mix} = \sqrt{\frac{1}{m} \sum_{q=1}^{m} h_{jq}^2}, \tag{13}$$

$$h_{jq} = \|\mathbf{g}_j\| \sin \theta_j^q. \tag{14}$$

It can be found that the only difference between $\mathrm{Tr}(\mathbf{F}_{\mathrm{mc}}^{-1}(\mathbf{l}))$ and $\mathrm{Tr}(\mathbf{F}_{\mathrm{mix}}^{-1}(\mathbf{l}))$ is the weighted average squared distance among gradients, namely $\overline{h}_w^{mix}$ and $\overline{h}_w^{mc}$.

Because the relative positions between the mobile device and the jth, qth APs are fixed, the angles are equal for both the single channel RSS measurements and the mixed channel RSS measurements, i.e.

$$\sin \theta_{ij}^{pq} = \sin \theta_j^q, \quad i = 1, ..., c \quad p = 1, ..., c. \tag{15}$$

Thus, substituting h_{ij}^{pq} in (7) and h_{jq} in (13) with (6) and (14), respectively, the difference between the corresponding pair of $(\overline{h}_{ij}^{mc})^2$ and $(\overline{h}_j^{mix})^2$ is

$$\begin{aligned} &\frac{1}{c} \sum_{p=1}^{c} \|\mathbf{g}_{pq}\|^2 \sin^2 \theta_{ij}^{pq} - (\frac{1}{c} \sum_{p=1}^{c} \|\mathbf{g}_{pq}\|)^2 \sin^2 \theta_j^q \\ =&\frac{\sin^2 \theta_j^q}{c^2} \left((c-1) \sum_{p=1}^{c} \|\mathbf{g}_{pq}\|^2 - 2 \sum_{p<r} \|\mathbf{g}_{pq}\| \|\mathbf{g}_{rq}\| \right) \\ =&\frac{\sin^2 \theta_j^q}{c^2} \sum_{p<r}^{c} (\|\mathbf{g}_{pq}\| - \|\mathbf{g}_{rq}\|)^2 \geq 0, \end{aligned} \tag{16}$$

where the equality holds if and only if $c = 1$. Therefore, given $i = 1, ..., c$ and $j = 1, ..., m$, the following inequality holds

$$\overline{h}_{ij}^{mc} \geq \min(\overline{h}_{ij}^{mc}) > \overline{h}_j^{mix}, \tag{17}$$

$$\begin{aligned} \overline{h}_w^{mc} =& \sqrt{\sum_{j=1}^{m} \left(\frac{\sum_{i=1}^{c} (\|\mathbf{g}_{ij}\| \overline{h}_{ij}^{mc})^2}{\sum_k \sum_r \|\mathbf{g}_{kr}\|^2} \right)} \\ \geq& \sqrt{\sum_{j=1}^{m} \left(\frac{\sum_{i=1}^{c} \|\mathbf{g}_{ij}\|^2}{\sum_k \sum_r \|\mathbf{g}_{kr}\|^2} (\min(\overline{h}_{ij}^{mc}))^2 \right)} \\ >& \sqrt{\sum_{j=1}^{m} \left(\frac{(\sum_{i=1}^{c} \|\mathbf{g}_{ij}\|)^2}{\sum_r (\sum_{k=1}^{c} \|\mathbf{g}_{kr}\|)^2} \overline{h}_j^{mix} \right)} = \overline{h}_w^{mix}. \end{aligned} \tag{18}$$

In lieu of the above analysis, we can obtain that, if $c > 1$,

$$\mathrm{Tr}(\mathbf{F}_{\mathrm{mc}}^{-1}(\mathbf{l})) < \mathrm{Tr}(\mathbf{F}_{\mathrm{mix}}^{-1}(\mathbf{l})), \tag{19}$$

which reveals that the localization error in the first case is less than that in the second case, implying that correctly utilizing multi-channel RSS measurements will benefit localization.

C. A Practical Issue

The above analysis assumes that, during online localization, c RSS measurements are respectively obtained from c channels by every AP, which is hard to be satisfied in practice: 1) neither commercial nor customized APs used in this paper are able to simultaneously sniff all the WiFi channels. 2) there exist collisions during packets transmissions, especially in the popular Channel 1, 6 and 11, such that some Probe request packets might be missed by APs. Therefore, straightforwardly extending the size of a fingerprint by c is not feasible. To address this issue, we adopt the following two solutions.

Firstly, we introduce channel sets to extend a traditional fingerprint, namely that several neighboring channels are combined into one channel set, such that the mean RSS measurement cross the channels in each channel set is evaluated in the new fingerprint. As a result, the fingerprint is extended by less than c times (assume n_c channel sets in total, $n_c < c$), but it also follows from the above theoretical analysis that the resulting localization error should still be mitigated.

Secondly, the usage of channel sets does not often result in sufficient RSS measurements for online localization [38], and as such, we adopt a sliding time window with a duration of Δt, so as to collect as many RSS measurements as possible. As a result, we can obtain the following measurement vector

$$\mathbf{r}^{\mathbf{\Delta t}} = \left[r_{11}^{\Delta t}, r_{12}^{\Delta t}, ..., r_{1m}^{\Delta t}, r_{21}^{\Delta t}, ..., r_{ij}^{\Delta t}, ..., r_{n_c m}^{\Delta t} \right], \tag{20}$$

where $r_{ij}^{\Delta t}$ denotes the mean RSS measurement in the ith channel set at the jth AP within Δt.

By taking into account RSS measurements from multiple channel sets, more useful features are leveraged to differentiate one position from another. Besides, this treatment can be directly applied in either existing localization methods or future ones, including the DHCLoc proposed in this paper.

IV. Optimal Offset Estimation for Heterogeneous Devices

This section introduces the offset variable to combat device heterogeneity that severely degrades localization performance.

To ease the presentation, we reformulate the online RSS measurement vector $\mathbf{r}^{\Delta t}$ in (20) into $\mathbf{r} = [r_1, ..., r_i, ..., r_n]$, where $n = n_c \cdot m$. The mean and standard deviation of fingerprints, which is obtained by using a RD, at the jth RP are denoted by $\mathbf{R}_j = [R_{1j}, ..., R_{ij}, ..., R_{nj}]$ and $\mathbf{\Sigma}_j = [\sigma_{1j}, ..., \sigma_{ij}, ..., \sigma_{nj}]$, respectively.

Given a target area $\mathcal{A}$, the unknown position $\mathbf{l}$ in $\mathcal{A}$ can be estimated by using MLE, i.e.

$$\hat{\mathbf{l}} = \arg\max_{\mathbf{l} \in \mathcal{A}} \log \prod_{i=1}^{n} p(r_i | \mathbf{l}). \tag{21}$$

It should be noted that, device heterogeneity comes from the usage of different chips, antennas and etc., and mainly leads to different transmission powers or receiver gains. As a result, there exists an offset between the RSS measurements from two heterogeneous devices at the same position, as observed in [20]. We intend to find out such an offset between any given device and the RD. To this end, we have to address

the following two issues: 1) the offset varies from device to device; 2) we do not have prior information on the device being used, so that we do not know the corresponding offset.

Define an individual offset for a specific device and introduce this variable into the MLE-based localization framework in (21). As such, the localization problem becomes finding the fingerprint most similar to the online RSS measurement vector calibrated by the offset, namely

$$\begin{aligned}\hat{\mathbf{l}} &= \arg\max_{\mathbf{l}\in\mathcal{A},a_j} \log\prod_{i=1}^{n} p(r_i|\mathbf{l},a_j)\\ &= \arg\max_{j,a_j} \log\prod_{i=1}^{n} p(R_{ij}=r_i-a_j|l_j),\end{aligned} \tag{22}$$

where a_j denotes the unknown offset. According to the Gaussian assumption on the fingerprint, we can have

$$\hat{\mathbf{l}} = \arg\max_{j}\arg\max_{a_j}\sum_{i=1}^{n}\left\{-\frac{1}{2\sigma_{ij}^2}[a_j^2+2(R_{ij}-r_i)a_j+(R_{ij}-r_i)^2]-\log\sigma_{ij}\right\}. \tag{23}$$

It can be clearly seen that, given a fixed j, the cost function (23) is quadratic with one unknown variable a_j, so that the solution, i.e., OOE, can be deduced by

$$a_j^{opt} = -\frac{\sum_i \frac{R_{ij}-r_i}{\sigma_{ij}^2}}{2\sum_i \frac{1}{2\sigma_{ij}^2}}. \tag{24}$$

Thus, (23) can be simplified as follows

$$\begin{aligned}\hat{\mathbf{l}} &= \arg\max_{\mathbf{l}\in\mathcal{A}} \log\prod_{i=1}^{n} p(r_i-a_j^{opt}|\mathbf{l})\\ &= \arg\max_{j} \frac{4\left(\sum_i \frac{1}{2\sigma_{ij}^2}\right)\sum_i\left(\frac{(R_{ij}-r_i)^2}{2\sigma_{ij}^2}+\log\sigma_{ij}\right)-\left(\sum_i\frac{R_{ij}-r_i}{\sigma_{ij}^2}\right)^2}{4\sum_i\frac{1}{2\sigma_{ij}^2}}.\end{aligned} \tag{25}$$

According to the above formulas, a_j^{opt} associated with each device can be directly calculated using the existing FD in the online stage, then the original RSS measurement vector can be calibrated by subtracting a_j^{opt} to cater for heterogeneity, and finally the RP with the highest probability is selected.

Specifically, OOE can be used in most existing localization algorithm after simple adjustments.

- For deterministic approaches using KNN, the calibrated fingerprint $\mathbf{r}' = \{[r_i - a^{opt}]|i\in\{1,...,n\}\}$ instead of the original one is used at each RP to calculate the Euclidean distance $dis = \sum_{i=1}^{n}(R_{ij}-r_i')^2$.
- For probabilistic approaches using WKNN, (25) can be directly used to calculate probabilities (weights).
- For supervised machine learning methods including DHCLoc, adversarial training [39] is adopted to generate robust models for accurately localizing heterogeneous devices, which will be described in the following section.

V. Proposed DHCLoc Algorithm

In this section, after presenting the DNN-based localization model, we summarize the whole process of DHCLoc.

A. DNN-based Localization Model

The DNN-based localization problem can be formulated as finding the optimal function, denoted $f^*(\mathbf{r}^{\Delta t})$, mapping a given RSS measurement matrix to a position, i.e.

$$\begin{aligned}\hat{\mathbf{l}} &= f^*(\mathbf{r}^{\Delta t}),\\ f^* \in \mathcal{F} &= \{f(\mathbf{r}^{\Delta t};\theta)|\theta\in\mathbb{R}^D\},\end{aligned} \tag{26}$$

where $\mathcal{F}$ is the set of possible mapping functions, θ is the hyper-parameters and D is the number of hyper-parameters. From the perspective of machine learning, utilizing RSS measurements from multiple channel sets is equivalent to extending the features of inputting RSS measurements. Therefore, we design the following DNN-based localization model.

1) Basic Architecture: The architecture of the DNN model is shown in Fig. 2, and its details are listed as follows.

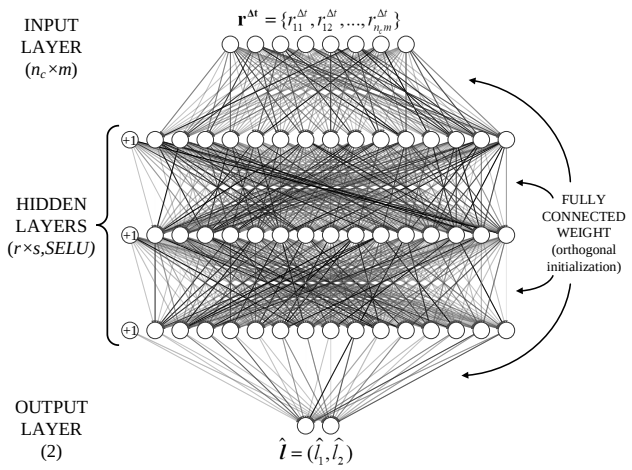


Fig. 2. The basic architecture of the proposed DNN model (thicknesses of connections reflect different weights).

- **Input Layer:** Take each $n_c\times m$ multi-channel RSS vector $\mathbf{r}^{\Delta t}$ as input to the corresponding input layer of DNN.
- **Architecture:** r layers and s units in each layer are specified as the fully connected hidden layers to extract features and predict positions.
- **Parameter Initialization:** The orthogonal initialization technology [40] is adopted to initialize the weights between each node pair in different layers to avoid gradient disappearance or explosion at the beginning of training.
- **Output Layer:** Because classification returns only a limited number of discrete outputs and often leads to large coordinate estimation errors, we adopt the multi-objective regression with 2 outputs corresponding to a pair of position coordinates.
- **Activation Function:** As for the activation function of each node, we employ the recently reported scaled exponential linear unit (SELU) [41], i.e.

$$selu(x) = \lambda\begin{cases} x, & if \quad x>0,\\ \alpha e^x-\alpha, & if \quad x\le 0,\end{cases} \tag{27}$$

where $\lambda\approx 1.0507$ and $\alpha\approx 1.6733$.

- **Loss Function:** We adopt the MSE as the overall loss function $\mathbb{E}_{(\mathbf{r}^{\Delta t},\mathbf{l})}$ of the proposed DNN model, namely

$$\mathbb{E}_{(\mathbf{r}^{\Delta t},\mathbf{l})} = MSE(\mathbf{l}) = \frac{1}{|\mathcal{D}|}\sum_{i=1}^{|\mathcal{D}|}(\mathbf{l}_i - \hat{\mathbf{l}}_i)^2, \tag{28}$$

  where $|\mathcal{D}|$ is the number of training samples, and $\mathbf{l}_i$ and $\hat{\mathbf{l}}_i$ are the real and predicted positions of the ith training sample.
- **Optimizer:** The newer adaptive momentum method for stochastic optimization [42] is used as the optimizer.
- **Early Stop:** To avoid overfitting as well as speed up the training process, early stop strategy [43] is adopted.

Obviously, according to the rationale of DNN model, we can easily obtain the time complexity of our model's online predicting process: $O(n_c \cdot m \cdot s + (r-1)\cdot s^2 + 2\cdot s + r\cdot s) = O(n^3)$ (let $n = \max(n_c \cdot m, r, s)$), and thus the time complexity of offline training process: $O(n^3 \cdot N)$ (where N denotes the iteration number of training). Therefore, once the limited application context is fixed, the overhead of the proposed model would stay the same magnitude with existing methods.

2) Data Augmentation for Multi-channel RSS: In order to 1) provide sufficient data to train DNN and 2) balance the uneven training samples cross different RSS channel sets, we leverage GPR [30] to generate a certain amount of training samples at each RP for every channel set and every AP. Assuming $r_{ij}^{\Delta t} \sim N(R_{ij}, \sigma_{ij}^2)$, and then re-sample each $r_{ij}^{\Delta t}(i = 1, 2, ..., n_c; j = 1, 2, ..., m)$ according to its probability density function (PDF) to construct $\mathbf{r}^{\mathbf{\Delta t}}$ as one item of the training samples.

3) Adversarial Training for Combating Device Heterogeneity: Given a fixed reference device, since the offset varies cross different devices in a bounded range, we assume that the optimal offset follows the following uniform distribution

$$a^{opt} \sim U(a^{min}, a^{max}). \tag{29}$$

Due to the fact that there is no a prior information about all heterogeneous devices, it is hard to accurately determine the parameters of the uniform distribution. Hence, we consider to use the range of a small number of unlabeled test devices' OOE statistics to approximate. The subtle difference between the approximate and actual distribution can be further tolerated by adversarial training, which can be extended to arbitrary environments with trivial costs. Finally, the process of adversarial training is summarized as follows

$$\min_{\theta} \mathbb{E}_{(\mathbf{r}^{\Delta t} + a^{opt}\mathbf{I_n},\mathbf{l})} \sim \mathcal{D}\left[L(f_\theta(\mathbf{r}^{\Delta t} + a^{opt}\mathbf{I_n}), \mathbf{l})\right], \tag{30}$$

where $\mathbf{r}^{\Delta t}$ is the original measurement vector, f_θ is the model function and L is the loss for a single adversarial sample. To be applicable in practice, the following steps are considered: 1) obtain the (multi-channel) RSS measurement vectors; 2) determine the limited number of test devices' positions by (25), and then select the nearest RPs and apply (24) to calculate the posteriori OOEs; 3) construct the uniform distribution (29) of a^{opt} using all OOEs; 4) generate adversarial samples with $\mathbf{r}^{\Delta t} + a^{opt}\mathbf{I_n}$; 5) train the model with all adversarial samples. With the newly generated adversarial samples, the optimization procedure of (30) will confuse the model as much as possible to enhance the tolerance of heterogeneous devices.

B. DHCLoc Algorithm

In summary, the offline training algorithm and online localization algorithm of the proposed DHCLoc are listed in Algorithm 1 and Algorithm 2. Based on the traditional localization framework, the algorithms incorporate multi-channel information (Section III), leverage adversarial training to combat device heterogeneity (Section IV), and then use the DNN model (Section V-A) to capture the features of multi-channel and adversarial samples.

Algorithm 1 Offline training algorithm of DHCLoc

1: **Original Data Collecting:** place the RD at each RP to collect original training data for a certain period (e.g., 1 min), and record [AP No., RP No., channel No., RSS measurement];
2: **FD Generating:** for each RP, select the RSS measurements at ith channel set and jth AP, calculate the means $\bar{r}_{ij}^{k}$ and standard deviation σ_{ij}^{k}, and obtain the fingerprint matrix of kth RP $\bar{\mathbf{r}}^{\mathbf{k}}$ and $\sigma^{\mathbf{k}}$. All the fingerprint matrices with its position coordinates are combined into the FD;
3: **Training Data Generating:** for each RP in FD, use GPR to re-sampling a certain amount of $\bar{\mathbf{r}}^{\mathbf{k}}$ as training data, such the size of training data is multiple times of FD;
4: **Adversarial Samples Constructing:** use statistics on a^{opt} of several unlabeled test devices to determine a^{max} and a^{min}, and then add noises to each training item by $(\mathbf{r} + a^{opt}\mathbf{I_n})$ with $a^{opt} \sim U(a^{min}, a^{max})$ and coordinates due to inaccurate position calibration when collecting original data [21];
5: **DNN Training:** use the preprocessed training data to train the proposed DNN model with max iteration N and early-stopping parameter κ;
6: **Hyperparameters Adjusting and Model Saving:** try the positioning effects of different hyperparameters (r and s) on the verification data (partitioning from training data), and save the best model on the hard disk.

Algorithm 2 Online localization algorithm of DHCLoc

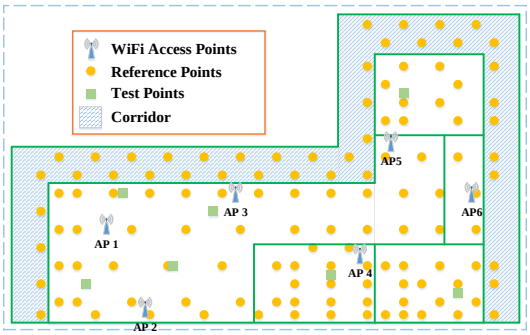
1: **Model Loading:** load the DNN model into the RAM of localization server;
2: **Time Window Data Partitioning:** intercept the real-time positioning data stream (or cut the testing data) with sliding time window of length Δt and step δ;
3: **Location Fingerprint Generating:** transform the intercepted data into location fingerprint $\bar{\mathbf{r}}^{\mathbf{\Delta t}}$ (analogy to $\bar{\mathbf{r}}^{\mathbf{k}}$);
4: **Localizing:** take $\bar{\mathbf{r}}^{\mathbf{\Delta t}}$ as an input of the DNN model, and use the output $\hat{\mathbf{l}} = (\hat{x}, \hat{y})$ as a location estimate.

VI. Experiments

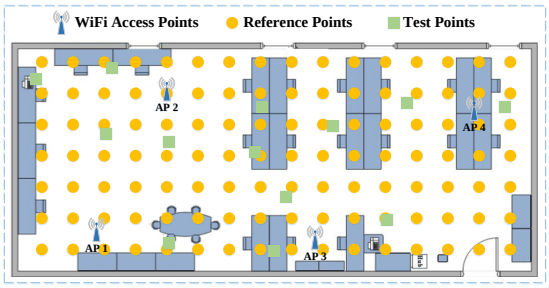
In order to validate our work, we carry out extensive experiments with two different datasets which were collected during several weeks using heterogeneous devices.

A. Setup

1) Testbeds: The experiments were conducted in two typical environments, termed OFFICE and LAB respectively. Their layouts are shown in Fig. 3 and the details are listed in Table I. The OFFICE testbed is a common indoor office scenario with the area of around $350m^2$, whereas the LAB testbed is a rectangular space with the area of $81m^2$. In each testbed, 6 and 4 customized WiFi APs (i.e. sniffers) were respectively installed, and each AP can simultaneously monitor 9 channels in both $2.4G$ and $5G$ bands [44]; 6 different mobile devices were used, as listed in Table I(b).



(a) OFFICE Testbed



(b) LAB Testbed

Fig. 3. The physical layouts of two testbeds.

2) Data Collection: The built-in program in each AP is able to periodically upload sniffing data, including RSS measurements, channel number, MAC address, timestamps, etc., to a server through UDP.

3) Channel Sets: As mentioned in Section III, RSS measurements from multiple channels are grouped by channel sets instead of single channel, so that four different schemes are considered as listed in Table II. Specifically, regarding the $5G$ band, RSS measurements are collected in only 4 channels, and considering the large bandwidth spanned, we only combine them into at most 2 channel sets.

4) Training and Testing Samples: In the OFFICE (LAB) testbed, 122 (112) grid points with the grid size of around $2m$ ($0.8m$) are assigned as RPs as shown in Fig. 3, and 100 RSS measurements were generated regarding each channel set and one AP. As for the testing data, 7 (13) positions were randomly assigned as testing points (TPs), and a suitable positioning interval $\delta = 1s$ and time window $\Delta t = 5s$ were adopted. Consequently, the RSS measurements between one TP and one AP within each interval are averaged for localization with respect to each channel set, i.e. $\bar{\mathbf{r}}^{\mathbf{\Delta t}}$.

5) Minimum Detected APs Number (MDAN): In practice, due to packet collisions and limited transmission coverage, the number of APs detecting an identical mobile device during a short period of time can vary from 1 to the number of total APs. As such, we impose a precondition of localization, namely that only if one mobile device is detected by no less than MDAN APs in every channel set, it will be considered for localization. The effect of MDAN on localization shall be investigated. In the experiments, we let MDAN be 3, and provided that a mobile device is not detected by one AP in a channel set, the corresonding RSS measurement is defautly set to be -100 dBm.

6) Baseline Methods for Comparison: The deterministic approach based on KNN, probabilistic approach based on WKNN, and the pure DNN model are employed as baseline methods. Given an online RSS measurement vector $\mathbf{r}^{\mathbf{\Delta t}}$, KNN selects k RPs with the least Euclidean distances between their fingerprints and $\mathbf{r}^{\mathbf{\Delta t}}$, and calculates the average of the coordinates of the k RPs as the final location estimate; WKNN selects k RPs at which $\mathbf{r}^{\mathbf{\Delta t}}$ attains the largest probabilities by assuming a Gaussian joint distribution for each fingerprint, and calculates the weighted average of the corresponding k pairs' of coordinates according to their probabilities. The existing popular device heterogeneity handling methods, i.e. SSD [19] and PA-based fingerprint standardization [20], are also implemented. DHCLoc and DNN are constructed using the Keras frame [45] in PyCharm, and the other methods are programmed with MATLAB R2017b.

B. Validating the Effectiveness of Using Multi-channel RSSs

Three baseline methods are implemented with respect to the four different channel set schemes. Both the experiments in [19], [20], [46] and our pilot studies indicate that KNN performs well with small k and WKNN with large k, and thus $k = 15$ for KNN and $k = 30$ for WKNN. The maximal iteration number of DNN is 1000 with the patience parameter of early-stopping being 5. Given different channel set schemes, the DNN model performs well in our fine-tuning studies when the layer number r and node number s lie in $3 \sim 6$ and $16 \sim 128$, respectively, suggesting that we should assign such values to the corresponding parameters. Consequently, the Cumulative Distribution Function (CDF) of resulting localization errors for each method is plotted in Fig. 4 with the $RMSE$ evaluated as follows

$$RMSE(f^*(\mathbf{RSS}^{\Delta t})) = \frac{\sum_{(\mathbf{RSS}^{\Delta t},\mathbf{l})\in\mathcal{D}'} \|f^*(\mathbf{RSS}^{\Delta t}) - \mathbf{l}\|}{|\mathcal{D}'|}, \quad (31)$$

where $|\mathcal{D}'|$ is the number of testing samples, $f^*(\mathbf{RSS}^{\Delta t})$ is localization result and $\mathbf{l}$ is the corresponding real position. Moreover, due to slight fluctuations caused by random functions in DNN, the $RMSE$ of DNN is evaluated by averaging over 10 times, which is a more fair comparison way compared to evaluating by the minimum $RMSE$.

As can be seen, regarding the KNN-based method and DNN model, using multi-channel RSS measurements always benefits localization accuracy, whereas the WKNN-based method shows worse performance in few cases (e.g., 4CS for OFFICE

TABLE I
THE DETAILS OF TWO TESTBEDS

(a) The parameters of two testbeds

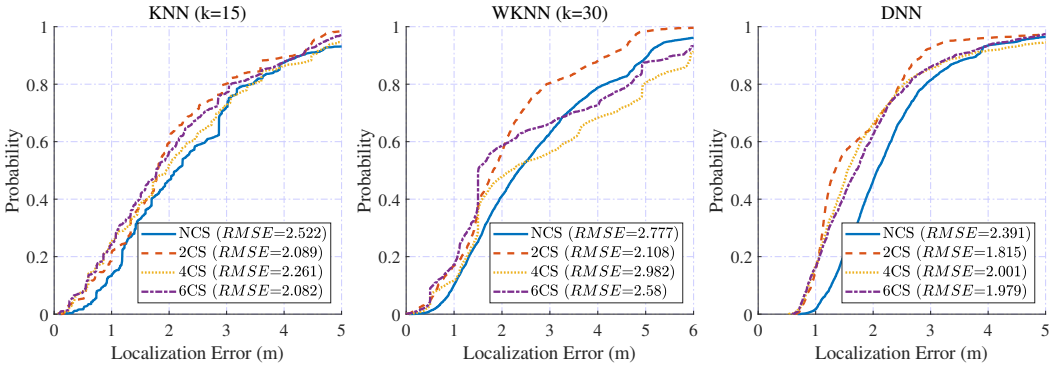
Dataset	Size	Shape	RP	TP	Devices
OFFICE	Medium ($18m \times 10m + 10m \times 17m$)	irregular	122	7	6 smartphones
LAB	Small ($13.5m \times 6m$)	rectangle	112	13	2 laptops&3 smartphones&1 tablet

(b) The details of mobile devices being used

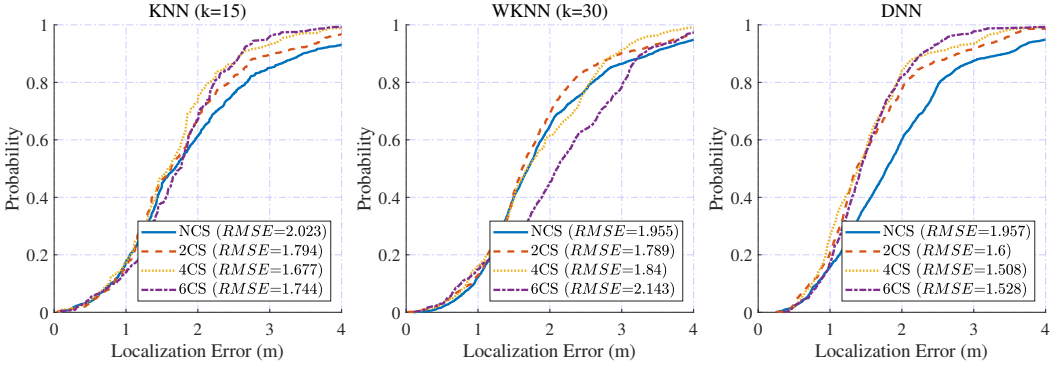
Dataset	Model	Short Name	2.4G Band	5G Band	Reference Device	Test Device
OFFICE	iPhone 7	iPhone	✓	✓	×	✓
	Huawei Honor 8X	8X	✓	✓	✓	✓
	Smartisan T2	T2	✓	✓	×	✓
	Huawei Mate9	Mate9-1	✓	✓	×	✓
	Huawei Mate9	Mate9-2	✓	✓	×	✓
	Huawei P8	P8	✓	×	✓	✓
LAB	Xiaomi M6	Mi6-1	✓	✓	✓	✓
	Xiaomi M6	Mi6-2	✓	✓	✓	✓
	Dell Laptop	Dell PC	✓	×	×	✓
	Sansung S9+	S9+	✓	✓	×	✓
	Macbook Pro Laptop	Macbook	✓	✓	×	✓
	iPad Pro Tablet	iPad	✓	✓	×	✓

TABLE II
FOUR DIFFERENT CHANNEL SETS

Scheme Name	Short Name	Channel Merging	Used in
non–channel splitting	NCS	{1, 2, 3, 6, 8, 9, 11, 13, 42, 58, 155, 165}	Section VI-B,VI-C,VI-D
2 channel sets	2CS	{1, 2, 3, 6, 8, 9, 11, 13}, {42, 58, 155, 165}	Section VI-B,VI-D
4 channel sets	4CS	{1, 2, 3, 6}, {8, 9, 11, 13}, {42, 58}, {155, 165}	Section VI-B,VI-D
6 channel sets	6CS	{1, 2, 3}, {6}, {8,9}, {11, 13}, {42, 58}, {155, 165}	Section VI-B



(a) Results on the OFFICE dataset



(b) Results on the LAB dataset

Fig. 4. The CDF of localization errors with respect to four different channel set schemes.

TABLE III
THE *RMSE*S OF HETEROGENOUS DEVICES EVALUATED BY THREE BASELINES WITH DIFFERENT HETEROGENEITY HANDLING APPROACHES

(a) Results on the OFFICE dataset

Methods	**Devices**						**Average**
$k = 15, wk = 30$	iPhone	8X (RD)	T2	Mate9-1	Mate9-2	P8 (RD)	
KNN	2.2	2.761	**2.299**	2.46	3.079	**2.278**	2.513
SSD-KNN	2.173	2.615	2.583	**1.471**	2.826	2.71	2.396
PA-KNN	2.368	2.757	2.715	2.092	**2.773**	2.785	2.582
OOE-KNN	**2.151**	**2.611**	2.573	1.751	2.804	2.337	**2.371**
WKNN	3.938	3.192	**2.465**	2.138	4.153	**1.94**	2.971
SSD-WKNN	3.14	3.531	3.511	2.585	3.042	3.649	3.243
PA-WKNN	2.603	**2.959**	3.149	2.789	**2.777**	2.877	2.859
OOE-WKNN	**2.473**	3.151	3.005	**1.61**	3.015	2.682	**2.656**
DNN	2.42	2.664	**1.929**	1.747	2.88	2.354	2.332
SSD-DNN	2.098	2.531	2.327	**1.298**	2.518	2.42	2.199
PA-DNN	1.963	2.32	2.329	1.408	**2.401**	2.34	2.126
OOE-DNN	**1.951**	**2.227**	1.983	1.491	2.408	**2.185**	**2.041**

(b) Results on the LAB dataset

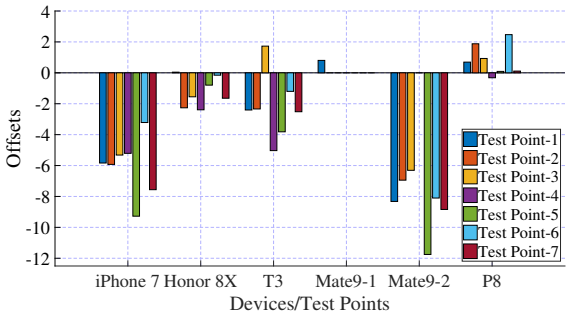
Methods	**Devices**						**Average**
$k = 15, wk = 30$	Mi6-1 (RD)	Mi6-2 (RD)	Dell PC	S9+	Macbook	iPad	
KNN	1.506	1.532	1.975	2.871	1.746	2.392	2.004
SSD-KNN	1.673	1.611	2.021	3.181	1.828	2.277	2.099
PA-KNN	1.749	1.56	2.073	2.782	1.722	**2.147**	2.006
OOE-KNN	**1.495**	**1.528**	**1.903**	**2.061**	**1.664**	2.225	**1.813**
WKNN	1.749	**1.623**	**1.955**	**2.016**	1.845	2.417	**1.934**
SSD-WKNN	1.933	2.002	2.328	2.83	2.145	2.541	2.297
PA-WKNN	1.816	1.698	2.134	2.745	**1.663**	**2.256**	2.052
OOE-WKNN	**1.738**	1.716	2.029	2.306	1.72	2.323	1.972
DNN	1.842	1.629	1.931	**2.262**	1.783	2.288	1.956
SSD-DNN	1.784	1.607	1.875	2.328	1.774	2.288	1.943
PA-DNN	**1.724**	1.608	1.863	2.578	**1.667**	**2.138**	1.93
OOE-DNN	1.78	**1.585**	**1.824**	2.277	1.729	2.241	**1.906**

and 6CS for LAB). Meanwhile, DNN always outperforms other alternatives, which confirms the powerful learning ability of the deep learning technique. However, localization accuracy does not always rise as increasing the number of channel sets, which is attributable to the fact that having more channel sets risks the lack of RSS measurements in more channel sets and consequently degrades localization accuracy.

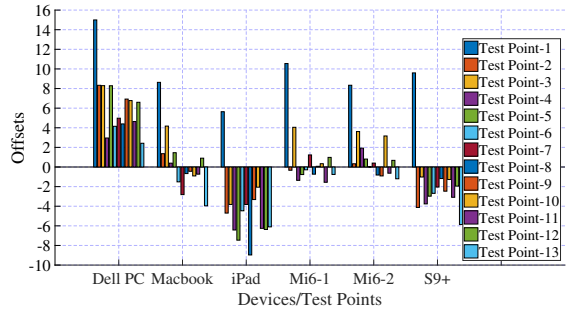
C. Validating the Effectiveness of OOE

Three baseline methods with and without using OOE, SSD and PA are compared. Regarding the DNN model, adversarial training is adopted to fuse OOE as described in Section V-A3. We designedly remove the location labels of all test devices so as to avoid using apriori information. By calculating the OOEs at TPs for OFFICE and LAB with NCS as depicted in Fig. 5, we can obtain $a^{opt}_{OFFICE} \sim U(-12, 3)$ and $a^{opt}_{LAB} \sim U(-9, 15)$, which are added into the original training data. Besides, all the other settings including parameters are the same as above.

The *RMSE*s are listed in Table III, and it can be concluded that: 1) due to the complexity and randomness of the real-world data, in few cases the *RMSE*s of the methods dealing with device heterogeneity can be slightly larger than those of traditional KNN, WKNN and DNN, but the methods based on OOE lead to the smallest average *RMSE*s, which validates the effectiveness of OOE; 2) as we have mentioned

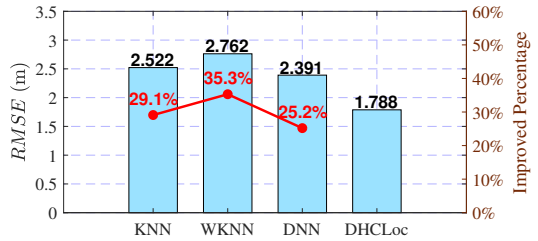


(a) OFFICE with NCS

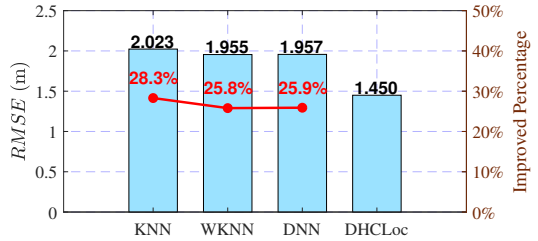


(b) LAB with NCS

Fig. 5. The results of OOE with the NCS scheme.

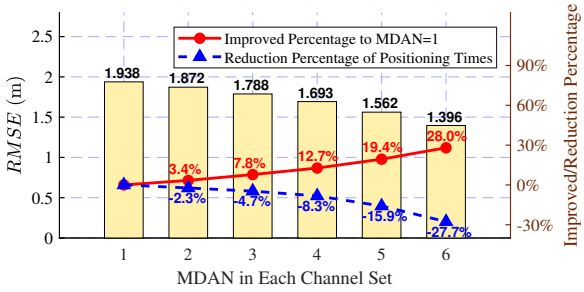


(a) Results on the OFFICE dataset

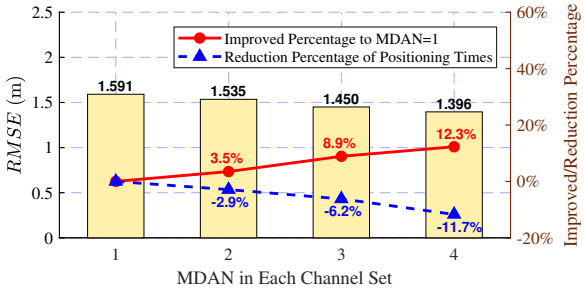


(b) Results on the LAB dataset

Fig. 6. The $RMSE$s of DHCLoc and baseline methods.



(a) Results on the OFFICE dataset with 2CS



(b) Results on LAB dataset with 4CS

Fig. 7. The $RMSE$ of DHCLoc with different MDANs in each channel set.

in Section II-B, the methods based on SSD and PA cannot achieve ideal performance in the situation with a relative small number of APs, whereas the methods based on OOE demonstrate evident superiority in both stability and accuracy.

D. Validating the Overall Performance of DHCLoc

According to the results in Section VI-B, DHCLoc adopts the 2CS scheme for OFFICE and the 4CS scheme for LAB, respectively. In addition, we can obtain $a^{opt}_{OFFICE(2cs)} \sim U(-5,9)$ and $a^{opt}_{LAB(4cs)} \sim U(-8,8)$. The $RMSE$s of four methods, and corresponding reduction percentages achieved by DHCLoc are plotted in Fig. 6. As can be seen, DHCLoc attains the best localization accuracy, and particularly, outperforms the baseline methods by at least 25.2% on OFFICE and 25.8% on LAB, respectively.

Furthermore, to investigate the effect of MDAN, the reduction percentages of localization times and the improved percentages of localization accuracies compared to the case with MDAN= 1, are plotted in Fig. 7. It can be seen that, with the MDAN in one channel set increasing, both the accuracy and the reduction of localization time will accordingly increase in a similar degree, but the difference slowly declines. This result inspires us that, we should reasonably select MDAN so as to balance the localization accuracy and frequency. In addition, it can be found that, under the strictest conditions, i.e. MDAN= 6 for OFFICE and MDAN= 4 for LAB, both of the $RMSE$s on two datasets achieved by DHCLoc are less than $1.4m$, which shows a great potential on more fine-grained WiFi fingerprinting localization.

VII. Conclusion

In this paper, we proposed a DNN-based WiFi localization method, namely DHCLoc, for passive WiFi localization. Specifically, channel set splitting and adversarial training are employed to conquer the two challenges arising from multi-channel RSS measurements and device heterogeneity, which were validated by both theoretical analyses and experiments. The experiments showed the overall performance of DHCLoc outperforms several existing algorithms by at least 25.2% and 25.8% with two field datasets respectively, indicating that DHCLoc is superior to the state-of-the-art methods by a significant margin.

Regarding future works, we will focus on new advanced DNN models, such as CNN and RNN, and the adaptability of DHCLoc in different environments through transfer learning.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No. 41871363, 41761086 and 61761035), the Key Science-Technology Project of Inner Mongolia (Grant No. 2021GG0163) and the Natural Science Foundation of Inner Mongolia Autonomous Region (Grant No. 2021ZD13 and 2019MS06030).

References

[1] L. Hao, B. Huang, H. Hong, B. Jia, and W. Li, "A channel adaptive wifi indoor localization method based on deep learning," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, 2021, pp. 1–6.

[2] B. Huang, L. Xie, and Z. Yang, "Tdoa-based source localization with distance-dependent noises," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 468–480, 2015.

[3] B. Badihi, J. Zhao, S. Zhuang, O. Seppänen, and R. Jäntti, "Intelligent construction site: On low cost automated indoor localization using bluetooth low energy beacons," in *2019 IEEE Conference on Wireless Sensors (ICWiSe)*, 2019, pp. 29–35.

[4] J. Tian, Z. Zhou, J. Wu, S. Du, C. Xiang, and C. Kuang, "Synthetization of fingerprint recognition and trilateration for wi-fi indoor localization through linear kalman filtering," in *China Satellite Navigation Conference (CSNC) 2018 Proceedings*. Singapore: Springer Singapore, 2018, pp. 373–386.

[5] X. Wang, L. Gao, S. Mao, and S. Pandey, "Csi-based fingerprinting for indoor localization: A deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 763–776, Jan 2017.

[6] S. He and S. . G. Chan, "Wi-fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.

[7] Q. Jiang, K. Li, M. Zhou, Z. Tian, and M. Xiang, "Competitive agglomeration based knn in indoor wlan localization environment," in *2015 10th International Conference on Communications and Networking in China (ChinaCom)*, Aug 2015, pp. 338–342.

[8] L. Yen, C. Yan, S. Renu, A. Belay, H. Lin, and Y. Ye, "A modified wknn indoor wi-fi localization method with differential coordinates," in *2017 International Conference on Applied System Innovation (ICASI)*, 2017, pp. 1822–1824.

[9] J. Rahman, N. Mondal, K. Islam, and M. A. Hasan, "Feature fusion based svm classifier for protein subcellular localization prediction," *Journal of Integrative Bioinformatics*, vol. 13, 03 2016.

[10] M. Abbas, M. Elhamshary, H. Rizk, M. Torki, and M. Youssef, "Wideep: Wifi-based accurate and robust indoor localization system using deep learning," in *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*, 2019, pp. 1–10.

[11] A. Belay, H.-P. Lin, G. Tarekegn, and S.-S. Jeng, "Applying deep neural network (dnn) for robust indoor localization in multi-building environment," *Applied Sciences*, vol. 8, p. 1062, 06 2018.

[12] G. Xiong, T. Kim, and E. Perrins, "Decorrelation deep learning for fingerprint-based indoor localization," 2019.

[13] M. Ibrahim, M. Torki, and M. ElNainay, "Cnn based indoor localization using rss time-series," in *2018 IEEE Symposium on Computers and Communications (ISCC)*. Los Alamitos, CA, USA: IEEE Computer Society, jun 2018, pp. 548–1049.

[14] X. Song, X. Fan, C. Xiang, Q. Ye, L. Liu, Z. Wang, X. He, N. Yang, and G. Fang, "A novel convolutional neural network based indoor localization framework with wifi fingerprinting," *IEEE Access*, vol. 7, pp. 110 698–110 709, 2019.

[15] L. Deng and D. Yu, "Deep learning: Methods and applications," *Foundations and Trends in Signal Processing*, vol. 7, 01 2013.

[16] B. Huang, G. Mao, Y. Qin, and Y. Wei, "Pedestrian flow estimation through passive wifi sensing," *IEEE Transactions on Mobile Computing*, vol. 20, no. 4, pp. 1529–1542, 2021.

[17] B. Huang, M. Liu, Z. Xu, and B. Jia, "On the performance analysis of wifi based localization," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018, pp. 4369–4373.

[18] H. Cheng, F. Wang, R. Tao, H. Luo, and F. Zhao, "Clustering algorithms research for device-clustering localization," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2012, pp. 1–7.

[19] A. K. M. Mahtab Hossain, Y. Jin, W. Soh, and H. N. Van, "Ssd: A robust rf location fingerprint addressing mobile devices' heterogeneity," *IEEE Transactions on Mobile Computing*, vol. 12, no. 1, pp. 65–77, 2013.

[20] H. Zou, B. Huang, X. Lu, H. Jiang, and L. Xie, "A robust indoor positioning system based on the procrustes analysis and weighted extreme learning machine," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1252–1266, 2016.

[21] Y. Wei and R. Zheng, "Handling device heterogeneity in wi-fi based indoor positioning systems," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2020, pp. 556–561.

[22] P. Bahl and V. Padmanabhan, "Radar: An in-building rf-based user location and tracking system," vol. 2, 02 2000, pp. 775 – 784 vol.2.

[23] S. He and S.-H. G. Chan, "Wi-fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.

[24] C.-t. Huang, C.-h. Wu, Y.-n. Lee, and J.-t. Chen, "A novel indoor rss-based position location algorithm using factor graphs," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 3050–3058, 2009.

[25] Y. Jin, W.-S. Soh, and W.-C. Wong, "Indoor localization with channel impulse response based fingerprint and nonparametric regression," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 1120–1127, 2010.

[26] H. Zou, L. Xie, Q.-S. Jia, and H. Wang, "An integrative weighted path loss and extreme learning machine approach to rfid based indoor positioning," in *International Conference on Indoor Positioning and Indoor Navigation*, 2013, pp. 1–5.

[27] Y. Shu, Y. Huang, J. Zhang, P. Coué, P. Cheng, J. Chen, and K. G. Shin, "Gradient-based fingerprinting for indoor localization and tracking," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2424–2433, 2016.

[28] C. Wu, J. Xu, Z. Yang, N. D. Lane, and Z. Yin, "Gain without pain: Accurate wifi-based localization using fingerprint spatial gradient," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 2, Jun. 2017.

[29] J. Xu, Z. Yang, H. Chen, Y. Liu, X. Zhou, J. Li, and N. Lane, "Embracing spatial awareness for reliable wifi-based indoor location systems," in *2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. Los Alamitos, CA, USA: IEEE Computer Society, oct 2018, pp. 281–289.

[30] B. Huang, Z. Xu, B. Jia, and G. Mao, "An online radio map update scheme for wifi fingerprint-based localization," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6909–6918, Aug 2019.

[31] J. Park, D. Curtis, S. Teller, and J. Ledlie, "Implications of device diversity for organic localization," in *2011 Proceedings IEEE INFOCOM*, 2011, pp. 3182–3190.

[32] C. Figuera, J. L. Rojo-Álvarez, I. Mora-Jiménez, A. Guerrero-Curieses, M. Wilby, and J. Ramos-López, "Time-space sampling and mobile device calibration for wifi indoor location systems," *IEEE Transactions on Mobile Computing*, vol. 10, no. 7, pp. 913–926, 2011.

[33] A. W. Tsui, Y.-H. Chuang, and H.-H. Chu, "Unsupervised learning for solving rss hardware variance problem in wifi localization," *Mob. Netw. Appl.*, vol. 14, no. 5, p. 677–691, Oct. 2009.

[34] F. Della Rosa, H. Leppäkoski, S. Biancullo, and J. Nurmi, "Ad-hoc networks aiding indoor calibrations of heterogeneous devices for fingerprinting applications," in *2010 International Conference on Indoor Positioning and Indoor Navigation*, 2010, pp. 1–6.

[35] D. Munoz, F. Bouchereau, C. Vargas, and R. Enriquez, in *Position Location Techniques and Applications*. Oxford: Academic Press, 2009.

[36] T. Rappaport, *Wireless Communications : Principles and Practice, 2e*. Prentice Hall, December 2001.

[37] Y. Wen, X. Tian, X. Wang, and S. Lu, "Fundamental limits of rss fingerprinting based indoor localization," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 2479–2487.

[38] B. Huang, R. Yang, B. Jia, W. Li, and G. Mao, "A theoretical analysis on sampling size in wifi fingerprint-based localization," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3599–3608, 2021.

[39] F. Tramèr, A. Kurakin, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel, "Ensemble adversarial training: Attacks and defenses," 2017.

[40] A. M. Saxe, J. L. McClelland, and S. Ganguli, "Exact solutions to the nonlinear dynamics of learning in deep linear neural networks," 2013.

[41] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," 06 2017.

[42] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 12 2014.

[43] S. Scardapane, M. Scarpiniti, E. Baccarelli, and A. Uncini, "Why should we add early exits to neural networks?" *Cognitive Computation*, Jun 2020.

[44] "Shenzhen daison intelligence technology co. ltd." http://www.daison-intelligence.com/, 2018.

[45] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[46] H. Zou, B. Huang, X. Lu, H. Jiang, and L. Xie, "Standardizing location fingerprints across heterogeneous mobile devices for indoor localization," in *2016 IEEE Wireless Communications and Networking Conference*, 2016, pp. 1–6.