OHetTLAL: An Online Transfer Learning Method for Fingerprint-Based Indoor Positioning

Hailu Tesfay Gidey 1, Xiansheng Guo 1,2,*, Ke Zhong 1, Lin Li 1 and Yukun Zhang 1

1 Department of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
2 Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou 324000, China
* Correspondence: xsguo@uestc.edu.cn; Tel.: +86-132-5810-3244

Abstract: In an indoor positioning system (IPS), transfer learning (TL) methods are commonly used to predict the location of mobile devices under the assumption that all training instances of the target domain are given in advance. However, this assumption has been criticized for its shortcomings in dealing with the problem of signal distribution variations, especially in a dynamic indoor environment. The reasons are: collecting a sufficient number of training instances is costly, the training instances may arrive online, the feature spaces of the target and source domains may be different, and negative knowledge may be transferred in the case of a redundant source domain. In this work, we proposed an online heterogeneous transfer learning (OHetTLAL) algorithm for IPS-based RSS fingerprinting to improve the positioning performance in the target domain by fusing both source and target domain knowledge. The source domain was refined based on the target domain to avoid negative knowledge transfer. The co-occurrence measure of the feature spaces (\(C_{mi}^{n}\)) was used to derive the homogeneous new feature spaces, and the features with higher weight values were selected for training the classifier because they could positively affect the location prediction of the target. Thus, the objective function was minimized over the new feature spaces. Extensive experiments were conducted on two real-world scenarios of datasets, and the predictive power of the different modeling techniques were evaluated for predicting the location of a mobile device. The results have revealed that the proposed algorithm outperforms the state-of-the-art methods for fingerprint-based indoor positioning and is found robust to changing environments. Moreover, the proposed algorithm is not only resilient to fluctuating environments but also mitigates the model’s overfitting problem.

Keywords: indoor positioning; online transfer learning; machine learning algorithms; heterogeneous feature spaces; optimization

1. Introduction

With the rapid proliferation of sensor-equipped mobile devices, intelligent applications, and the ubiquity of wireless communication networks, as well as the growing demand for location-based services (LBS), localization, tracking, and navigation systems are receiving much attention from researchers, engineers, practitioners, and the industry [1–3]. Location services have greatly facilitated people’s daily activities because the services they need are mostly location-based. Although Global Positioning System (GPS) is the standard method for location determination and is known for its reliable and promising accuracy outdoors, it cannot provide accurate location determination indoors due to non-line-of-sight (NLOS) propagation, poor penetration of GPS signals in complex buildings, and the multipath effect caused by walls and obstacles [4–6]. On the other hand, rapid commercial growth and user demand for location-based services and social networking services (SNSs) have made Indoor Positioning Systems (IPSs) essential in all areas of life where mobility plays a role [1,7,8]. Moreover, unlike
traditional mobile services, LBSs are used in mobile applications where the user’s current location is needed to provide context-aware functionality [9]. As a result, IPSs are becoming an important research area and are considered a critical element in pervasive computing [10,11]. Indoor positioning is a system for locating mobile devices or objects in a complex indoor environment where GPS signals are attenuated or blocked [34].

Various indoor positioning technologies have been proposed in the literature, such as radio-frequency technologies (RFID) [12], Bluetooth (BLE) [13], ultra-wideband (UWB) [14], Zigbee [15], inertial navigation system [16], Visible Light Communication (VLC) [17], etc., to address the complex nature of the indoor environment, although high implementation costs are required for the additional infrastructure. In the last two decades, various classifications for indoor location technologies have been used [18,19]. In [20], indoor location technologies are classified into building-dependent and building-independent technologies based on the sensors used for localization. According to [20], building-dependent technologies include Wi-Fi, cellular, and Bluetooth, which use the building’s infrastructure, and technologies such as RFID, UWB, infrared, ultrasound, Zigbee, VLC, and acoustic signals, which require their infrastructure. In contrast, image-based technologies and dead reckoning are classified as building-independent. Indoor positioning systems based on wireless networks for building communities, especially Wi-Fi and BLE as they may leverage the building’s infrastructure, have gained significant acceptance from researchers and industry due to their low cost and ease of installation [21–27]. Along with this, precise information is generally required in modern building domain for indoor localization with the use of BLE network technology including but not limited to emergency management [21], smart plug load control [22], smart HVAC controls [23], route guidance [24], activity recognition [25], monitor occupancy [26], epidemic prevention and health monitoring [27], and others.

A study in [27], proposes machine learning approaches for a smart contact tracing (SCT) BLE-based system with the goal of epidemic prevention and control. The system used each smartphone’s RSS measurements extracted from the BLE network to estimate its distance from other user’s phones and an alert is issued when social distancing rules are violated. Moreover, the decision tree algorithm has been found to detect potential cases to infectious individuals with an accuracy of 90%, given that two users hold their smartphone in a similar manner. In [21], they have also extracted information from beacons installed in a building which communicate with an occupant’s mobile application for estimating the occupant’s location in an indoor environment (considered both the commercial setting and indoor emergency situation) using machine learning algorithms. In addition, their results [21] revealed that combining BLE with machine learning is certainly promising as the basis for occupancy estimation. It has been also noted that BLE technology has significantly lower power consumption and much lower bit rates than Wi-Fi technology and is ideal for regular short-distance data transmission [28]. Thus, BLE technology is being used or limited for a short length of the advertising packet. Furthermore, privacy security protocols need to be preserved to ensure that the BLE’s beacon packet broadcasting will not reveal one’s identity, otherwise scalability of the technology will be negatively affected.

The two well-known indoor localization approaches are the (a) fingerprint-based approach and (b) range-based approach. The range-based approach requires accurate estimates of geometric parameters based on signal features such as received signal strength (RSS) [29], arrival times (TOA) [30], angle of arrival (AOA) [31], and time differences of arrival (TDOA) [32] for the position of the target to obtain a better position estimate. Due to the complex nature of the indoor environment, where radio signals are characterized by the following: NLOS propagation, multipath effects, and a dynamic environment, the distance-based approach cannot provide accurate geometric parameters [4–6]. On the other hand, the fingerprint-based approach has attracted much attention in indoor localization because it is independent of geometric parameter estimation and provides better results than the distance-based approach in a complex indoor environment [4–6]. Thus, Wi-Fi RSS-based indoor fingerprint positioning (FPBIP) could be implemented without additional infrastructure, as it is widely used for communication purposes [33,34].
and is very important for both terminal devices and sensor networks in Internet-of-Things (IoT) applications [35,36]. Therefore, we used Wi-Fi RSS-based fingerprinting, the most popular indoor localization technique [1,33–36], and included two stages: Offline (training) and Online (localization). In the offline stage, fingerprints are collected from different locations, and the predictive model is trained to learn the relationship between signal and location. The learned model is then used to determine the location of the mobile device or user based on the observed new measurements in the online phase [34–37].

However, the positioning performance of the FPBIP approach is not robust and efficient due to the dynamic indoor environment. Researchers have proposed several methods to deal with the dynamic indoor environment, which leads to low localization accuracy due to variations in fingerprint patterns over time. These methods can be classified into four groups: probabilistic methods [38,39], machine learning methods [40,41], exploiting the quality of fingerprints of various signal features (such as Channel State Information (CSI) [42], Phase of Arrival (PoA) [43], Time of Flight (ToF) [44], etc.), and derived signal features (Signal Strength Difference (SSD) [15], and Fused Group of Fingerprints [45,46]). Although the proposed methods [38–46] have improved the location accuracy, they also suffer from the fluctuations of the signal distribution and are not robust to indoor dynamic environments. This is mainly because (i) the measurements depend on individual fingerprints, (ii) the fingerprint database needs to be up to date, which is impractical, and (iii) a sufficient number of labeled samples are required, which is also very costly. In parallel, hybrid location systems [8,47] (HPS) have been proposed to solve the single location problem, and the results demonstrate better location performance than the single system. The results show better positioning performance than that of a single system. This could be due to the fact that the combination of the two technologies has a leverage effect that compensates for each other’s errors. However, a hybrid base station falls outside the scope of our work and is not economically feasible. On the other hand, computational complexity is a serious problem for hybrid systems based on indoor positioning.

Therefore, it is important to develop a method that reduces the offline calibration effort and still achieves relatively high location accuracy. Transfer learning (TL) methods are effective for many real-world applications to leverage knowledge from the labeled training data of a source domain to improve model performance in a target domain for which there is little or no labeled target training data [48]. However, most TL methods assume that the source and target domains follow the same distribution, so their applications [49] are limited to indoor positioning. Moreover, the labeled training data of the target domain is limited, and it is a difficult task to construct a precise correspondence map to connect the source and target domains. Therefore, we proposed a novel technique to connect the source and target domains to utilize the knowledge of the source and target domains to improve the positioning performance in the target domain. The contributions in this paper are as follows:

1. We proposed a novel method of construction of a refined source domain (SD) through the feature space transformation technique to enhance quality attributes of the SD in the sense that to avoid negative transfer knowledge to the target domain (TD).

2. We build a fusing weighting approach to exploit both offline and online instances to enhance the learning performance in the target domain through transfer learning.

3. Effectiveness of the proposed approach was measured with different machine learning algorithms considering as baselines and verified by conducting extensive experiments on real-world datasets.

The remaining of the paper is organized as follows: Related works are presented in Section 2. Section 3 described fingerprinting localization framework and its problem formulation. Experimental results, discussions and evaluation metrics are presented in Section 4. Conclusions are provided in Section 5.
2. Related Works

There are two main parts to this section: the current state of indoor positioning systems, and the proposed framework of online heterogeneous transfer learning (OHetTL) applied to the indoor positioning problem (IPP).

2.1. Indoor Positioning System

Recently, due to the increased demand for LBS and applications, several indoor localization methods have been proposed and implemented [1–3,37], but they are not suitable for indoor localization due to some technical limitations and additional infrastructure investment costs [33–35,50]. As in [20] indicated, building-dependent technologies include Wi-Fi, cellular, and Bluetooth, which use the building’s infrastructure, and technologies such as RFID, UWB, infrared, ultrasound, Zigbee, VLC, and acoustic signals, which require their infrastructure. In contrast, image-based technologies and dead reckoning are classified as building-independent. Due to their low cost and ease of installation, indoor positioning systems based on wireless networks for building communities, particularly Wi-Fi and BLE as they may utilize the building’s infrastructure, have attracted a lot of attention from researchers and the industry [21–27]. In [22], they have also described the feasibility of BLE deployment in addressing the plug load management systems to reduce the rising energy consumption of plug loads in commercial buildings through different load monitoring and control strategies. They have further discussed a novel IoT-based occupancy-driven plug load management system called Plug-Mate, which mainly relies on the users’: (a) high-resolution occupancy information obtained through a nonintrusive indoor localization system, (b) plug load type information inferred through an advanced plug load identification feature, and (c) diverse control preferences through a personalized user interface. On the other hand, the authors described the fact that deployment locations of the BLE occupancy sensors are highly dependent on two main determinants: (a) a constant power availability to operate the sensors and (b) communication range of the BLE signal, causing sensors to be set 10–15 m apart from each other and evenly distributed to ensure complete signal coverage [22].

Another study in [51] proposed a scalable and less intrusive occupancy detection method-based on smartphones’ device using BLE network technology to perform zone-level occupancy localization, without the need for a mobile application. The results [51] revealed that despite the supervised ensemble model producing the best performance in terms of accuracy and macro average f1-score, the semi supervised clustering model demonstrated practical advantages, as it was able to produce a reasonable performance, while using a fraction of the training data. This could justify that the semi supervised clustering algorithms might be effective in the occupant’s detection for limited labeled samples. Moreover, a non-intrusive occupancy monitoring approach, which leverages on existing BLE technologies found in smartphone devices, has been proposed to track the occupants’ movement patterns using BLE beacons [26]. The proposed approach does not require the installation of a mobile application, just the occupant’s MAC address of their Bluetooth-enabled smartphone devices [26]. Generally, BLE technology has significantly lower power consumption and much lower bit rates than Wi-Fi technology, and is ideal for regular short-distance data transmission [28]. Thus, BLE technology is being used or limited for a short length of advertising packet. In addition to that, intrusive indoor positioning-based BLE technology using the user’s mobile device are becoming more challenging for the scalability or feasibility of deployment due the concern for the user’s privacy. Thus, a privacy security protocol needs be preserved to ensure that the BLE’s beacon packet broadcasting will not reveal one’s identity, otherwise scalability of the technology will be negatively affected. This is consistent with the claims that have been discussed in [52]; regardless, indoor positioning-based BLE network technology is gaining importance due to its ubiquitous nature, low cost, and flexibility, but further scientific developments are limited for novel technologies such as BLE, due to the lack of open
datasets and corresponding frameworks suitable to compare and evaluate specialized localization solutions.

However, Wi-Fi fingerprinting [33–37] is becoming one of the most commonly used methods for indoor localization due to the ubiquity of Wi-Fi infrastructure and the popularity of Wi-Fi-enabled mobile devices. An IPS for commercial buildings has been developed to provide fine-grained occupancy-based HVAC (Heating, Ventilation and Air Conditioning systems (HVAC)) actuation [23]. This system exploits the existing Wi-Fi network infrastructure and occupants’ smartphones with the Wi-Fi enabled. They have also discussed about how using the existing infrastructure for occupancy detection greatly lowers the cost and labor of implementation and maintenance. Furthermore, noisy Wi-Fi signals and metadata regarding the building’s occupants, access points, and HVAC zone have also been employed to alleviate the difficulties in occupancy sensing, which could make it a cost-effective technique. Another method of Wi-Fi indoor localization based on SVM is proposed [53] to identify the floor using the Wi-Fi signal of each floor, and the method achieved an accuracy rate of 99.09% for 3D positioning [53]. An indoor localization system-based Wi-Fi fingerprint using spatial multi-points matching has also been proposed to estimate the user’s position [54]. An indoor location tracking technology-based Wi-Fi fingerprint technique [55] has been proposed to improve the user’s location accuracy using mobile communication technology and location tracking technology. Additionally, for an indoor IoT application using a Bayesian network and a limited radio map, a reliable 3D indoor positioning system-based Wi-Fi RSS fingerprinting has been developed [56]. However, to counteract the impact of received signal variations caused by multipath and signal attenuations throughout various time periods, the radio map should be updated within short time frame.

This method does not use radio signal propagation geometry but requires data acquisition and a built-in radio map in the offline phase, although it has notable shortcomings in the dynamic indoor environment and requires enough labeled samples, making it labor-intensive and costly [57]. Moreover, relatively satisfactory accuracy can be achieved by this method, but [33–35] the database fingerprint must be up to date, otherwise, the location performance will be severely degraded. Some research works have been conducted to reduce the effort and time required to create the radio map. Most notably, these include crowd-sourcing [58] and simultaneous localization and mapping over Wi-Fi [59]. However, crowd-sourcing requires the active participation of users or achieves low accuracy, and the Wi-Fi simultaneous localization and mapping approach also suffers from the high computational cost. In [60], it was also described how to speed up computation and enhance system accuracy for indoor environments by using a clustering approach based on fingerprinting. Compared to the non-clustering approach, which had an average distance error of 3.4 m, the suggested technique obtained an average distance error of 2.4 m. Various methods have been proposed in the literature to address the dynamic indoor environment, which results in low localization accuracy due to variations in fingerprint patterns over time. These methods can be classified into four groups: probabilistic methods [38,39], machine learning methods [40,41], using the quality of fingerprints of various signal features (such as Channel State Information (CSI) [3,42], Phase of Arrival (PoA) [43], Time of Flight (ToF) [44], etc.), and derived signal features (Signal Strength Difference (SSD) [15] and fused group of fingerprints [45,46]).

Although the proposed methods [15,38–46] have improved the location accuracy, they also suffer from fluctuations in the signal distribution and are not robust to a dynamic indoor environment, mainly due to two limitations: (i) the measurements depend on a single fingerprint and cannot represent the dynamic scenario of the indoor environment; (ii) the need for a sufficient number of labeled samples is also very costly. Therefore, a hybrid positioning system (HPS) has been proposed for various seamless localization applications (Wi-Fi, Bluetooth, UWB, and ZigBee integrated into a hybrid base station (BS)) to solve the standalone positioning problem [8]. Other researchers in [47] have also proposed a hybrid indoor localization system based on the IMU sensor and smartphone camera,
which has better positioning performance than the standalone system due to the possible combination of errors compensated by each technology. However, hybrid BS is not the subject of our work and is not economically feasible. On the other hand, computational complexity is a serious problem for hybrid indoor positioning systems. Research-based on machine learning algorithms (ML) have also been proposed to address the indoor location problem [40,41,61,62]. A limited number of researchers have focused on reducing the calibration overhead for indoor WLAN locations [63,64]. The impact of reducing the number of samples collected at each coverage point, and reducing the number of coverages points on the accuracy of location determination has been studied [63,64]. To compensate for the performance degradation with reduced calibration, both linear and kernel-based interpolation methods have been proposed [63,64] to patch an incomplete radio map. The hidden Markov model [64] and hierarchical Bayesian model [65] have also been proposed to exploit the information contained in the unlabeled user traces and improve the accuracy of location estimation. A label propagation method [66,67] has also been used to learn from labeled and unlabeled data under the assumption that similar data samples should have similar labels. However, this approach is not suitable for online prediction.

2.2. Online Transfer Learning Approach to Localization Problem

Several machine learning approaches [61,62] have achieved promising results assuming that the training and testing data have the same distribution. However, in practice, it is not possible to collect enough training data that belong to the same feature space as the test data due to resource and environment constraints. Transfer learning (TL) therefore addresses such cross-domain learning problems by extracting useful information from data in a related domain and transferring it for use in the target tasks. We believe that leveraging the knowledge from the labeled source data can drastically reduce the expensive data labeling effort. Depending on the feature spaces of the source and target domains, TL is divided into two categories: homogeneous transfer (HomTL) and heterogeneous transfer (HetTL). HomTL is used in many real-world applications where the source and target data follow the same distribution, such as text search [68], image classification [69], and face recognition [70]. Although the focus of this work is on IPS, a multiclassification problem that uses binary scores in a source domain has been proposed to alleviate the problem of sparse numerical scores in the target domain [71].

Another study highlighting HetTL for multiclass problems was performed based on the compressed sensing theory [72]. In addition, co-occurrence data was used to solve OHetTL problems [73–76]. A scheme was also explored to evaluate the relatedness between given source domains by transferred weights learned from co-occurrence data [77]. We also found that most existing studies on TL operate in offline batch learning environments, where it is assumed that all training instances of the target domain are given in advance [78,79]. However, this assumption does not hold in many real-world scenarios where collecting sufficient training data is very costly and training instances may arrive online. In the online transfer learning (OTL) problem, we aim to perform an online learning task in a target domain by leveraging knowledge from offline source data and sequentially updating a classifier based on feedback from a data sequence by processing each instance as it arrives [80,81].

To address the dynamic indoor environmental conditions, several transfer learning algorithms (TLALs) have been proposed. For example, one study assumed that the RSS value at each location can be computed as a linear aggregate of the RSS values of neighboring benchmarks for each Wi-Fi access point and that this linear relationship remains the same over different time periods [82]. It was found that a model tree was created for each access point at each site. Accordingly, it transmits the Hidden Markov Model (HMM) for a new time point [83]. It uses the neighborhood relation as a bridge to transfer the knowledge and treats the trace as a Markov chain, then the HMM is trained for the new time period. It also considers the problem of cross-device transfer learning and treats multiple devices as multiple learning tasks [84]. However, these algorithms consider either cross-time or cross-device transfer learning and none of them are applied in a real indoor
wireless localization system (WILS). The OTL task has also been studied in the context of multi-armed bandits [85]. Moreover, OTL has been applied with online homogeneous source data [86,87]. For online homogeneous TL problems, an ensemble strategy was used, and for OHetTL learning problems, a multi-view approach was used. However, the studies [88,89] considered OHetTL under the assumption that the feature space of the source domain is a subset of the feature space of the target domain. Another study also used some offline data to support an online task in a target domain [88–90]. In this sense, the Wi-Fi-based indoor localization system aims to determine the location of the user or mobile device based on the signals received from different Wi-Fi access points (APs). In this work, the proposed model would learn based on the derived new feature spaces refined through the co-occurrence weight score of instances to predict the labeling in the target domain that could arrive in an online (OHetTL). Thus, the Wi-Fi APs send the signal and the mobile user receive the RSS values. The localization process conducted at the user’s side connected to a server. Moreover, Figure 1 below explicitly describes the system operation used for RSS-based Fingerprinting of Indoor Positioning System (FBIPS).

![System Operation for RSS-based Fingerprinting of Indoor Positioning System](image)

**Figure 1.** System Operation for RSS-based Fingerprinting of Indoor Positioning System.
3. Problem Formulation and Framework

Consider an indoor environment can be divided into $R$ reference points (RP), which represent the target’s location, and each RP is numbered with a label $l_i (i = 1, 2, \ldots, L)$. Suppose there are $P$ detectable Wi-Fi access points $(j)$ in total, and the received signal strengths (RSS) of fingerprints at the $i$th reference point can be represented as a matrix: $M_i = [r_{1i}(i), r_{2i}(i), \ldots, r_{Pi}(i)]^T$ and represents the $i$th RSS values collected from the $i$th reference point of the $j$th access point $(i = 1, 2, \ldots, n_i; j = 1, 2, \ldots, p)$. In the OHetTL problem, the target instances $D_T = D_{oT} = \{X_T\} = \{(x^n_{mT})\}_{m=1}^{n_T} \in X$ arrive in an online manner, and the labeled source data $D_{ol} = \{X_{ol}, Y_{ol}\} = \{(x_i, y_i)\}_{i=1}^{n_{ol}} \in X \times y$, come from an offline source domain, where $n_T$ and $n_{ol}$ represent the numbers of target and source data, respectively. In the traditional machine learning (ML) techniques, both the training and testing instances be assumed to come from the same domain, and follow the same data distribution. In practice, however, training and testing data may come from different domains, which has either:

- Different marginal distributions, or different feature spaces: $X_S \neq X_T$, or $P(X_S) \neq P(X_T)$ or
- Different predictive distributions, or different label spaces: $y_S \neq y_T$, or $(P_S(y|x) \neq P_T(y|x))$.

Several machine learning algorithms have been proposed for indoor positioning problems, but they do not consider the possible effects of signal variations that are likely to originate from the source domain to improve the model’s positioning performance in the target domain. In practice, there may be multiple source domains associated with the target domain that could improve target positioning, although combining these domains is challenging (see Figure 2). The expensive data labeling effort can also be reduced by using the knowledge of the labeled source data. It is also much easier and cheaper to collect unlabeled samples in the field during the testing phase. Therefore, it is important to develop a method that reduces the effort required for offline calibration while still achieving relatively high accuracy in location determination. Figure 2 illustrates the approach of the proposed algorithm, which uses the knowledge from the source and target domains of heterogeneous feature spaces, considering the offline fingerprints of the radio map and the test fingerprints obtained during the online phase as the source and target domains, respectively. First, the proposed algorithm OHetTLAL would derive the new feature spaces based on the weighted co-occurrence of measurements of the offline and online fingerprints, and the newly derived feature spaces are called co-occurrence data. Hoping to improve the learning performance in the target domain, we then fused the offline data and the co-occurrence data to build the offline decision function. Finally, the positioning performance of the model in the target domain is improved by fusing both the offline and online decision functions. In addition, the loss function between the actual and estimated positions was evaluated to update the model. The $\oplus$ and $\otimes$ symbols used in Figure 2 represent the data fusion and the updating phase of the classifier based on the loss measurement, respectively. Figure 2 below illustrates the proposed framework of heterogenous transfer learning-based weight co-occurrence for the indoor positioning system.

3.1. Distribution Divergence Minimization

In real world scenarios, for the indoor positioning environment, the feature spaces of the sources domain and target domain are different for several practical reasons. Therefore, using the maximum mean discrepancy (MMD) criterion distributions of the two domains was compared, and we worked to minimize the distance between the sample mean of the source and target data (reducing the marginal distribution distance of the two domains). To calculate the marginal distribution distance between the source domain and the target domain, the following formula was adopted:

$$Dist(X_s, X_t) = \frac{1}{n_s} \sum_{i=1}^{n_s} x_s - \frac{1}{n_t} \sum_{m=1}^{n_t} x_t \parallel_2^2$$  \hspace{1cm} (1)
Figure 2. Proposed framework of heterogenous transfer learning-based weight co-occurrence for indoor positioning.

3.2. Derivation of the New Source Feature Spaces

There are two major problems with OHetTL: (a) the feature spaces of the source and target domains differ. In this manner, the offline database could not be used directly to enhance learning performance, as it could result in a negative knowledge transfer that would degrade positioning prediction accuracy. The labeled training data of the target domain are limited, making it difficult to build a precise correspondence map linking the source and target domains. Thus, our first step in addressing these challenges is to develop new feature spaces that achieve positive knowledge transfer for the target’s localization and are computationally efficient. In order to achieve this, we established a set of standards for constructing new feature spaces called the refined source domain (SD). Exclusion criteria were used to remove redundant knowledge from the offline database during the construction of the new source feature spaces. By excluding irrelevant RSS measurements from different Wi-Fi access points, extraneous features have no longer be included in the analysis. By analyzing the co-occurrence nature of the signal measurements or Wi-Fi access points, this procedure is not only about avoiding or minimizing all negative knowledge transfer to the target domain, but also about identifying the most significant predicting factors.

However, we believe that the similarity between Wi-Fi access points, which is the co-occurrence of the cross-domain Wi-Fi access points, could be used in place of scrutinizing the individual RSS measurements, as the coverage area of a detected access point may approximate the target’s position to some degree. The MMD criterion [91,92] distributions of Wi-Fi access points of the two domains was computed to compare the similarity of those access points (on average) and the motive for this access point similarity would indicate that the coverage area of a detected Wi-Fi access point on average could better approximate the target’s location. Suppose there are P Wi-Fi access points (Wi-Fi APs) detected in both offline and online phases. We define $J_{of}^{m,p} = [J_{of}^{1,p}, J_{of}^{2,p}, \ldots, J_{of}^{m,p}]^T$ and $J_{ol}^{m,p} = [J_{ol}^{1,p}, J_{ol}^{2,p}, \ldots, J_{ol}^{m,p}]^T$ to denote the set of detected Wi-Fi APs for offline and online phases, respectively, where $J_{of}^{m,p}, J_{ol}^{m,p} \in \{0, 1\}$. A binary response of the detection indicators would take a value of 1 when the corresponding Wi-Fi APs is detected, otherwise 0 for the undetected, for the $p$th access point in the $i$th offline instance and $m$th online instance, respectively. To construct a
would imply that the more detected Wi-Fi APs are shared by these two domains, and
as:
\[ X_j \cap X_n \]

dimensional features, we can select all elements in these specified dimensions to obtain the
\[ J \]

can be represented as
\[ D \]

respectively. In OHetTL, the source feature space
\[ C \]

have been refined based on the target domain to avoid negative knowledge transfer. The
co-occurrence measures of feature spaces (\( C_{mi}^p \)) were used to derive the homogeneous new
feature spaces, and those features with higher weight scores were selected to construct the
source domain and used as an input to train the classifier, as it could improve the
target’s location prediction. The pseudo code to construct the source domain is provided
in Algorithm 1.

Algorithm 1. Derivation of new source feature space

Input: (1) Offline sample \( D_{of} \); (2) Target domain \( D_T \); (3) The number of instances \( n_T \);
Output: Source domain \( D_S \)

1. \( D_S = [] \)
2. for \( i = 1, 2, \ldots, L \) do
3. \hspace{1em} for \( m = 1, 2, \ldots, n_T \) do
4. \hspace{2em} for \( i = 1, 2, \ldots, n_{of} \) do
5. \hspace{3em} Compute co-occurrence \( C(m, i) \) using Equation (2)
6. \hspace{1em} end for
7. \hspace{1em} Obtain the co-occurrence vector of the \( m \)th target sample
8. end for
9. end for
10. for \( i = 1, 2, \ldots, n_S \) do
11. \hspace{1em} Find \( i, m \ s.t \ C(m,i) = \max C_{mi} \)
12. \hspace{1em} \( x_{sp}^o = x_{SP}^o, y_{sp}^o = y_{SP}^o \)
13. \hspace{1em} \( D_S = \{ (x_{SP}^o, y_{SP}^o) \} \)
14. end for
15. return \( D_S \)

3.3. Construction of New Homogeneous Feature Spaces

Consider that there are \( P \) Wi-Fi APs detected during offline and online phases as
\[ J_{of}^p = \left[ j_{11}^p, j_{12}^p, \ldots, j_{1n_{of}}^p \right]^T \] \& \[ J_{ol}^m = \left[ j_{11}^m, j_{12}^m, \ldots, j_{1n_{ol}}^m \right]^T \]
respectively, where \( j_{1i}^p, j_{1i}^m \in \{0, 1\} \). Let the target RSS measurements arrive in an online manner given as:
\( D_T = \{ X_T \} = \{ (x_{yi})_{m=1}^{n_T} \in x \& D_{of} = \{ X_{of}, Y_{of} \} = \{ (x_{yi}, y_{yi}) \}_{i=1}^{n_{of}} \in x \times y \), for the offline database,
respectively. In OHetTL, the source feature space \( D_{of} = \mathbb{R}^{d_{of}} \) is different from the target
feature space \( D_T = \mathbb{R}^{d_{of}} \). The APs shared by the source domain and the target domain
can be represented as \( I_c = \left( \bigcup_{i=1}^{n_S} j_{1i}^p \right) \cap \left( \bigcup_{m=1}^{n_T} j_{1m}^m \right) \). Assuming that shared APs can offer dc-
dimensional features, we can select all elements in these specified dimensions to obtain the
source and target data with common features, denoted as:
\[ X_{S}^{\theta_{LC}} = \left[ X_1^{\theta_{LC}}, X_2^{\theta_{LC}}, \ldots, X_{S}^{\theta_{LC}} \right]^T \]
and
\[ X_{T}^{\eta_{TC}} = \left[ X_1^{\eta_{TC}}, X_2^{\eta_{TC}}, \ldots, X_{T}^{\eta_{TC}} \right]^T \]
respectively, where \( X_{S}^{\theta_{LC}} \) and
\[ X_{T}^{\eta_{TC}} \]
are in \( \mathbb{R}^d \). The specific Wi-Fi APs of the source and target domains can be given as:
\[ J_{S}^{(sp)} = \left( \bigcup_{i=1}^{n_S} j_{1i}^p \right) \backslash I_c \] and
\[ J_{T}^{(sp)} = \left( \bigcup_{m=1}^{n_T} j_{1m}^m \right) \backslash I_c \] respectively, where ‘\( \backslash \)’ denotes the set
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z \) respectively, where \( x_{sp} \) and \( x_{sp}^{m} \) are
assigned to the less related source samples. The weights are therefore updated as follows.
The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
To solve the above constrained optimization problem, Equation (3) can be converted into
an unconstrained optimization using the Lagrangian multiplier method. Let us assume
that we have obtained a location estimate \( z^{(l-1)} \) at the \((l-1)\)th iteration, and now we need

\[
\begin{align*}
\min & \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T,hm}^{mi} - x_{T,hm} \|_2^2 \\
\text{s.t.} & \sum_{i=1}^{n_S} e^{-w_{mi}} = 1, m = 1, 2, \ldots, n_T 
\end{align*}
\]

\[(3)\]

The weights constraint is to better minimize the signal measurements of differences
between the samples. The equality constraint mentioned in the objective function would
assign higher weights to the most related source samples and the lessor weights would be
assigned to the less related source samples. The weights are therefore updated as follows.
to estimate the weights $w(t)$ at the $t$th iteration. By applying the Lagrangian multiplier method, we obtain:

$$L(w_m, \alpha) = \sum_{m=1}^{n_T} \sum_{i=1}^{n_S} w_{mi} \| x_{T, hm}^m - x_{S, hm}^i \|_2^2 + \alpha \left( \sum_{i=1}^{n_S} e^{-w_{mi}} - 1 \right) = 0$$

(4)

where $\alpha$ is the Lagrangian multiplier. By letting the partial derivative of the Lagrangian with respect to $w_{im}$ and $\alpha$ to be zeros, we obtain:

$$\frac{\partial}{\partial w_{im}} L(w_m, \alpha) = \begin{cases} \| x_{T, hm}^m - x_{S, hm}^1 \|_2^2 - \alpha e^{-w_{1m}} = 0 \\ \| x_{T, hm}^m - x_{S, hm}^2 \|_2^2 - \alpha e^{-w_{2m}} = 0 \\ \vdots \\ \| x_{T, hm}^m - x_{S, hm}^{n_s} \|_2^2 - \alpha e^{-w_{nm}} = 0 \\ \sum_{i=1}^{n_s} e^{-w_{im}} - 1 = 0 \end{cases}$$

(5)

By adding up the first $n_s$ terms in Equation (5), we can obtain

$$\alpha = \sum_{i=1}^{n_s} \| x_{T, hm}^m - x_{S, hm}^i \|_2^2$$

(6)

Substituting Equation (6) into Equation (5) gives the estimated transfer weight as:

$$w_{mi} = -\ln \left( \frac{\| x_{T, hm}^m - x_{S, hm}^i \|_2^2}{\sum_{i=1}^{n_s} \| x_{T, hm}^m - x_{S, hm}^i \|_2^2} \right)$$

(7)

### 3.5. Prediction Evaluation Metrics

In this section, effectiveness of the proposed algorithm was measured with different machine learning algorithms taken as baselines and verified through extensive experiments based on real world datasets. The root mean square error (RMSE) was used to evaluate the effectiveness of the proposed algorithm and defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{m=1}^{n} \left( \hat{x}_m - x \right)^2 + \left( \hat{y}_m - y \right)^2}$$

(8)

where $[\hat{x}_m, \hat{y}_m]^T$ and $[x, y]^T$ are the predicted location estimate and the true location of a client of the 2-dimensional coordinates of the $m$th positioning sample, respectively. $n$ is the total number of samples to be located in the target domain.

### 4. Experimental Results and Discussion

In this section, two real-world experiments were carried out to evaluate our proposed algorithm. Experimental settings and datasets were first presented, and the overall performance of the classifiers are analyzed. Both feature spaces of scenarios were considered (before and after transformation).

#### 4.1. Experimental Settings

**4.1.1. Dataset A**

The experiment was conducted at UESTC (University of Electronic Science and Technology of China) as depicted in Figure 3 and the 21st floor of the Innovation building was taken as the indoor environment site consisting of 10 offices and one corridor covers an area of 1460 m$^2$ and partitioned into 210 reference points (RP), which represent the
target’s location, and each RP is numbered with a label \( l, (l = 1, 2, \ldots, L) \). \( n_{\text{of}} \) and \( n_{\text{id}} \) are the number of offline and online instances collected during the offline training phases and the online testing phase, respectively. The RSS of fingerprints at the \( l \)th RP can be represented as a matrix: \( M_l = [r_{l1}(i), r_{l2}(i), \ldots, r_{ln}(i)] \). Smartphones were used for the data collection at the deployment area.

### Figure 3. Experimental environments for dataset A showing the RP locations in blue circles.

#### 4.1.2. Dataset B

Another real-office experiment at UESTC (University of Electronic Science and Technology of China) on the 21st floor of the Innovation Building was also conducted in different environmental settings to validate our results, as shown in Figure 4. It covers an area of 1460 m\(^2\) and partitioned into 175 reference points (RP). Nine Wi-Fi access points were sparsely deployed to guarantee that at least three access points are detectable at each RP.

### Figure 4. Experimental layout conducted in UESTC for generating dataset B.

Moreover, Table 1 below provides the summary of the datasets used to evaluate the algorithms.

| Dataset | Metrics | Area (m\(^2\)) | UE          | Base Station (BS) | #BSs | #RPs |
|---------|---------|---------------|-------------|-------------------|------|------|
| A       | RSS     | 1460          | Smartphone  | APs               | 416  | 210  |
| B       | RSS     | 1460          | Smartphone  | APs               | 9    | 175  |

UE: User Equipment.
4.2. Distribution of Homogeneous Feature Spaces

4.2.1. Before Transformation

Tables 2 and 3 show the distributions of RSS measurements of the offline and online databases, respectively. There were 416 Wi-Fi access points and 15,660 instances were measured from each Wi-Fi access point. The mean RSS values for AP_0, AP_1, AP_2, AP_3, and AP_4 were −59.4, −64.9, −70.1, −74.2, and −73.3, respectively. A standard deviation of 16.8, 15.9, 14.5, 17.9, and 17.1 was also measured for the signal fluctuations. The mean RSS values for AP_412, AP_413, AP_414, and AP_415 are almost all −100, which means that the base stations are not detected at all (Table 2). The unique pattern for these Wi-Fi access points with a score of −100 is that the measured signal variations have a standard deviation of 0.11, 0.43, 0.15, and 0.27, while no signal was detected in any of these specified base stations. The lowest values of standard deviation observed in the above scenario suggest a higher consistency of measurements, even though the standard deviation (sd) can be surprisingly zero for some Wi-Fi access points. However, these base stations (with the lowest sd) were not detected on any mobile device. Such a value (−100) means that no base station was detected during the training and testing phase and we excluded it from the training model. Otherwise, it would be a negative knowledge transfer that either falsely inflate (both false positive and false negative predictions could be high) or degrade the target’s location prediction. Alternatively, the relationship between signal and location might not be characterized.

**Table 2.** The distribution and descriptive measures of the Offline RSS measures.

| SNo. | AP_0  | AP_1  | AP_2  | AP_3  | AP_4  | AP_412 | AP_413 | AP_414 | AP_415 |
|------|-------|-------|-------|-------|-------|--------|--------|--------|--------|
| Count| 15,660| 15,660| 15,660| 15,660| 15,660| 15,660 | 15,660 | 15,660 | 15,660 |
| Mean | −59.4 | −64.9 | −70.1 | −74.2 | −73.3 | −99.9  | −99.9  | −99.9  | −99.9  |
| std  | 16.8  | 15.9  | 14.5  | 17.9  | 17.1  | 0.11   | 0.43   | 0.15   | 0.27   |
| Min  | −100.0| −100.0| −100.0| −100.0| −100.0| −100.0 | −100.0 | −100.0 | −100.0 |
| 25%  | −73.0 | −76.0 | −81.0 | −90.0 | −86.0 | −100.0 | −100.0 | −100.0 | −100.0 |
| 50%  | −57.0 | −63.0 | −71.0 | −73.0 | −70.0 | −100.0 | −100.0 | −100.0 | −100.0 |
| 75%  | −45.0 | −52.0 | −58.0 | −58.0 | −59.0 | −100.0 | −100.0 | −100.0 | −100.0 |
| Max  | −24.0 | −30.0 | −36.0 | −40.0 | −36.0 | −92.0  | −93.0  | −94.0  | −92.0  |

**Table 3.** The distribution and descriptive measures of the Online RSS measures.

| SNo. | AP_0  | AP_1  | AP_2  | AP_3  | AP_4  | AP_412 | AP_413 | AP_414 | AP_415 |
|------|-------|-------|-------|-------|-------|--------|--------|--------|--------|
| Count| 4338  | 4338  | 4338  | 4338  | 4338  | 4338   | 4338   | 4338   | 4338   |
| Mean | −58.2 | −64.37| −76.53| −73.52| −100.0| −100.0 | −100.0 | −100.0 | −100.0 |
| std  | 15.63 | 15.14 | 14.40 | 17.92 | 0.00  | 0.00   | 0.00   | 0.00   | 0.00   |
| min  | −88.0 | −100.0| −100.0| −100.0| −100.0| −100.0 | −100.0 | −100.0 | −100.0 |
| 25%  | −73.0 | −78.0 | −89.0 | −88.0 | −100.0| −100.0 | −100.0 | −100.0 | −100.0 |
| 50%  | −56.0 | −63.0 | −77.0 | −72.0 | −100.0| −100.0 | −100.0 | −100.0 | −100.0 |
| 75%  | −45.0 | −52.0 | −65.0 | −58.0 | −100.0| −100.0 | −100.0 | −100.0 | −100.0 |
| max  | −24.0 | −32.0 | −38.0 | −40.0 | −100.0| −100.0 | −100.0 | −100.0 | −100.0 |

4.2.2. After Transformation

Tables 4 and 5 show the distributions of RSS measurements of the derived new features of the source domain and online databases, respectively. Based on the weighted co-occurrence factor, 41 feature spaces were selected to train the classifiers for further location prediction of the target. After creating the new feature spaces, a similar pattern was observed for the Wi-Fi Aps, which had a score of −100, and the signal fluctuations were also inflated to a standard deviation of 16.7, 15.8, 14.5, 17.9 and 14.5 for AP_0, AP_1, AP_2, AP_3, and AP_4, respectively. The higher values of standard deviation that were observed indicate less consistency in the measurements, although the standard deviation may be too large in relative terms for some Wi-Fi access points (with a higher weighting of
co-occurrence). Nevertheless, the base stations with a higher weighting of co-occurrence are the ones that have active signal responses to mobile devices in both phases.

Table 4. The distribution and descriptive measures of the Offline RSS measures.

| SNo. | AP_0 | AP_1 | AP_2 | AP_3 | AP_4 | AP_38 | AP_39 | AP_40 | AP_41 |
|------|------|------|------|------|------|-------|-------|-------|-------|
| Count | 15,660 | 15,660 | 15,660 | 15,660 | 15,660 | 15,660 | 15,660 | 15,660 | 15,660 |
| Mean | -59.3 | -64.9 | -70.1 | -74.2 | -81.1 | -98.9 | -98.9 | -95.32 | -94.2 |
| Std | 16.7 | 15.8 | 14.5 | 17.9 | 14.5 | 3.07 | 4.07 | 8.07 | 6.77 |
| Min | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 |
| 25% | -73.0 | -76.0 | -81.0 | -90.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 |
| 50% | -57.0 | -63.0 | -71.0 | -73.0 | -81.00 | -100.0 | -100.0 | -100.0 | -100.0 |
| 75% | -45.0 | -52.0 | -58.0 | -58.0 | -68.0 | -100.0 | -100.0 | -91.0 | -89.0 |
| Max | -24.0 | -30.0 | -36.0 | -40.0 | -50.0 | -83.0 | -68.0 | -66.0 | -69.0 |

Table 5. The distribution and descriptive measures of the Online RSS measures.

| SNo. | AP_0 | AP_1 | AP_2 | AP_3 | AP_4 | AP_38 | AP_39 | AP_40 | AP_41 |
|------|------|------|------|------|------|-------|-------|-------|-------|
| Count | 4338 | 4338 | 4338 | 4338 | 4338 | 4338 | 4338 | 4338 | 4338 |
| Mean | -58.2 | -64.3 | -76.5 | -73.5 | -80.40 | -98.9 | -98.3 | -100.0 | -91.4 |
| Std | 15.6 | 15.14 | 14.4 | 17.9 | 16.3 | 3.7 | 4.7 | 9.0 | 8.1 |
| Min | -88.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 |
| 25% | -73.0 | -78.0 | -89.0 | -88.0 | -100.0 | -100.0 | -100.0 | -100.0 | -100.0 |
| 50% | -56.0 | -63.0 | -77.0 | -72.0 | -82.00 | -100.0 | -100.0 | -91.0 | -100.0 |
| 75% | -45.0 | -52.0 | -65.0 | -58.0 | -65.0 | -100.0 | -100.0 | -84.0 | -86.0 |
| Max | -24.0 | -32.0 | -38.0 | -40.0 | -45.0 | -82.0 | -79.0 | -69.0 | -69.0 |

Table 6 shows that the number of measurements from each reference point is different and shows that the distribution is unbalanced for dataset A, so feature scaling was considered to avoid the effect of dominance of the larger occurrence of labels within the cluster. Otherwise, the larger features would dominate the others within the cluster. On the other side, the number of measurements collected from each RP demonstrates that the label distribution for dataset B is balanced. Thus, no feature scaling technique is required to avoid the dominance effect of the cluster’s higher label occurrence.

Table 6. Distributions of Wi-Fi received signal strength’s per label.

| Dataset A (#Labels = 210) | Labels | # of RSS values per label |
|--------------------------|--------|--------------------------|
|                          | 134    | 210                      |
|                          | 342    | 239                      |
|                          | 134    | 30                       |

| Dataset B (#Labels = 175) | Labels | # of RSS values per label |
|--------------------------|--------|--------------------------|
|                          | 127    | -                        |
|                          | 30     | 30                       |
4.3.1. Dataset A

The original number of feature spaces was 416 for dataset A. However, based on the co-occurrence weight scores and exclusion criteria, it was determined that several feature spaces were irrelevant to the analysis because no signal was detected in any of the specified access points. In addition, the Wi-Fi-received signal strength measurements that had the lowest co-occurrence weights for the offline and online source instances were excluded from the prediction task because they did not affect the localization estimate. Figure 5 shows that the RSS measurements for the different Wi-Fi access points have a certain pattern and follow different distributions for the different labels of Wi-Fi AP1 (AP_1). It mainly shows that the relationship between signal and location has been characterized and follows the unique pattern for a particular label. This shows that the signal variations of the received Wi-Fi signal strengths at a particular reference point have been measured with different distribution values, such as the standard deviation and the average values of the received signals. The unique pattern for the measurements collected by the Wi-Fi AP seems to have different values for the signal-to-location relationships of the different target points. In addition to that, there are some noticeable outlier values (as observed in Figure 5 for target points 1, 76, and 191) that may affect the distribution of the values or have a negative impact on the location accuracy (both false positive and false negative predictions could be high) or degrade the location prediction of the target point. Therefore, the relationship between signal and location could not be characterized, as well as the outliers could affect the location performance. Moreover, to better visualize the pattern of the signal distributions in a reference point, some portions of the labels have been extracted, as shown in Figure 6.

Figure 5. Distributions of RSS for Wi-Fi AP1 with their corresponding targets for dataset A.
Figure 6. Distributions of RSS for Wi-Fi AP1 with some portion of the targets for dataset A.

Figure 7 shows that the RSS measurements for the different Wi-Fi access points have a certain pattern and follow different distributions for the different labels of Wi-Fi AP4 (AP_4). It shows that the ‘signal-to-location’ relationship has been characterized and follows the unique pattern for the specific label. In other words, the signal variations of the Wi-Fi received signal strengths at a reference point were measured to have different distribution values, including the standard deviation and the average values of the received signals. The unique pattern for the measurements collected by the Wi-Fi AP also appears to have different values for the signal-to-location relationship of the different target points. Moreover, there are some noticeable outlier values (as shown in Figure 8 for target 1) that seem to have extremely fluctuating values at the time of measurement, which could affect the location determination of the target. Therefore, the relationship between signal and location could not be characterized because the outliers could affect the position estimation.

Figure 7. Distributions of RSS for Wi-Fi AP4 with their corresponding targets for dataset A.
Figure 8. Distributions of RSS from Wi-Fi AP4 for some portions of the targets of dataset A.

Figure 9 illustrates whether the Wi-Fi signal strength received at a grid point from multiple APs was assumed to be independent such that the Wi-Fi signals transmitted from different APs are transmitted independently and do not interfere with each other. Therefore, the distributions of the RSS measurements for pairs of Wi-Fi access points, especially after transformation techniques were being carried out, show a different pattern of fingerprints for a mobile user at a specific reference point. It shows that the RSS measurements have specific pattern for the different Wi-Fi APs of AP_1 and AP_2 and follows a different distribution of RSS values for the different labels of an access point (dissimilarity). Moreover, the ‘signal-to-location’ relationship was being characterized or studied. The distributions of the RSS values measured at a particular grid point are so dynamic in nature due to the inherent heterogeneity of indoor environment such that there is no guarantee that the signal fluctuations are to be represented by a single value at a specific position.

Figure 9. Distributions of RSS for Pair Wi-Fi APs of 1 and 2 for dataset A.
Figure 10 shows that the RSS measurements have a specific pattern for the different Wi-Fi APs of AP_7 and AP_8 and follows a different distribution of RSS values for the different labels of an access point. Moreover, the ‘signal-to-location’ relationship was being characterized, and the distributions of the RSS values measured at a particular grid point are dynamic in nature due to the inherent heterogeneity of the indoor environment, such that there is no guarantee that the signal fluctuations are represented by a single value at a specific position.

![Figure 10. Distributions of RSS for Pair Wi-Fi APs of 7 and 8 for dataset A.](image)

Figure 11a,b show the distributions of RSS measurements of the offline and online databases for dataset A, respectively. There were 416 Wi-Fi access points and 15,660 instances were measured from each Wi-Fi access point; the density distribution of the RSS measurements clearly demonstrate that the mean RSS values were observed to be $-100$, which means that instances were not received, as the majority of the Wi-Fi APs are not detected for the associated target points. The unique pattern for these Wi-Fi access points with an average score of $-100$ is that the measured signal variations have the lowest values of standard deviations, while no signal was detected in any of these specified APs. This is what the number tells, nevertheless, it is untrue. However, these APs (with the lowest standard deviation) were not detected on any mobile device during the training and testing phase, and we excluded it from the training model to avoid a negative knowledge transfer that could either falsely inflate (both false positive and false negative predictions could be high) or degrade the target’s position prediction. This is because we cannot talk about the signal distribution of reference points while the associated mobile devices could not receive any signal. Moreover, Figure 11c,d show that the distributions of RSS measurements of the derived new features of the source domain and online databases, respectively. Based on the weighted co-occurrence factor, a new refined source domain was constructed to train the classifiers for the position estimation of the target. One can observe at the density distributions of the refined source domain and target domain that the derived feature spaces are from different distributions with different values of parameters, hence the instances from the offline and online phases are heterogeneous.
Table 7 illustrates the predictive power of different predictive modeling techniques that have been used to evaluate and predict the target location of a mobile device based on the received signal strengths from different Wi-Fi APs. As shown in Table 6 for dataset A, the number of new feature spaces used for the analysis is 41 after the feature transformation was being carried out. In this case, the measurements coming from the remaining Wi-Fi access points were not included in the analysis because they could negatively affect the positioning task. This is because the co-occurrence weights of these Wi-Fi access points from both domains of the instances are insignificant, so including them would falsely inflate our positioning accuracy. Our experimental results demonstrate that the proposed algorithm (OHetTLAL) is the best algorithm with the lowest root mean square error (RMSE). This indicates that our proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from the related source domain to the target domain. Not only can a positive knowledge transfer from the source domain to the prediction task be observed, but also the computational cost and memory usage were effectively improved by excluding irrelevant features from the analysis. On the other hand, the Gaussian mixture model has achieved the lowest positioning accuracy and consumes a high computational cost.

![Figure 11. Distributions of RSS Wi-Fi APs of offline and online databases before and after transformation.](image-url)
Table 7. Models’ target prediction using the new feature spaces of 41 Wi-Fi access points based OHetTLAL.

| Classifiers                        | RMSE (in Meters) |
|------------------------------------|------------------|
| Decision tree                      | 1.72             |
| K - Neighbour (KNN)                | 1.69             |
| Support vector machine (SVC)       | 1.58             |
| Logistic Regression (LR)           | 2.41             |
| Random forest                      | 2.09             |
| Gaussian Mixture                   | 3.69             |
| Neural Network (MLP)               | 2.75             |
| The proposed algorithm (based TL)  | 1.08             |

Transfer Learning (TL) based.

Table 8 shows that experiments were also conducted for the indoor practical positioning dataset (Dataset A), and the position estimation accuracy of the different predictive modeling techniques was evaluated to train and estimate the location of a mobile device when Wi-Fi signal strengths were received from different Wi-Fi access points, with the dimensions of the new feature spaces reduced to 15 Wi-Fi access points. The measurements that came from the 26 Wi-Fi access points were also excluded from the analysis because the co-occurrence weight scores of these access points from both domains of the instances were insignificant, so including them would spuriously increase our position accuracy. Similarly, the proposed model was compared with seven machine learning algorithms (decision tree (DT), K nearest-neighbor (KNN), support vector machine (SVC), logistic regression (LR), random forest (RF), Gaussian Mixture Model (GMM), and neural network (MLP)) to see the impact after excluding irrelevant feature domains from the analysis based on their weight scores, hoping to avoid negative knowledge transfer. In addition, those measurements of Wi-Fi received signal strength that had the lowest co-occurrence weight values for the instances of offline and online sources were excluded from the prediction task because they had no impact on the positioning estimation. Moreover, insignificant and redundant features would further affect the computational cost and memory requirements of the algorithm. As shown in Table 8 for dataset A, the number of new feature spaces used for the analysis is 15. Thus, our experimental results demonstrate that the proposed algorithm is the best algorithm with the lowest root mean square error. On the other hand, the Gaussian mixture model (GMM) has obtained the highest error in positioning and consumes a huge computational cost. This indicates that our proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from the source domain to the target domain, and that transferring knowledge from the source domain to the prediction task positively affects the positioning accuracy.

Table 8. Models’ target prediction using the new feature spaces of 15 Wi-Fi access points based on OHetTLAL.

| Classifiers                        | RMSE (in Meters) |
|------------------------------------|------------------|
| Decision tree                      | 2.31             |
| K - Neighbour (KNN)                | 1.83             |
| Support vector machine (SVC)       | 1.79             |
| Logistic Regression (LR)           | 2.53             |
| Random forest                      | 2.52             |
| Gaussian Mixture                   | 4.24             |
| Neural Network (MLP)               | 1.97             |
| The proposed algorithm (based TL)  | 1.35             |

Figure 12 confirms that the proposed algorithm is the best algorithm with the lowest root mean squared error for dataset A with 41 feature spaces. In this case, the measurements collected from the remaining Wi-Fi access points were not included in the analysis because they could negatively affect the prediction task. The coincidence weights of these Wi-Fi
access points from both domains of instances were found to be insignificant, so including them would spuriously increase the positioning accuracy. All classifiers achieved minimal RMSE even after applying transfer learning. This indicates that our proposed algorithm succeeded in improving the positioning accuracy by transferring knowledge from the corresponding source domain to the target domain. It can also be observed that positive knowledge from the source domain was transferred to the prediction task. On the other hand, the Gaussian mixture model also achieved the lowest positioning accuracy.

Figure 12. Comparisons of classifier’s RMSE for dataset A with 41 Wi-Fi APs.

Figure 13 shows that the proposed algorithm was the best algorithm with the lowest root mean squared error for dataset A with 15 feature spaces. In this case, the measurements that came from the 26 Wi-Fi access points were also excluded from the analysis because the co-occurrence weight scores of these access points from both domains of the instances were insignificant, so including them would falsely inflate our positioning accuracy. Similarly, all classifiers achieved a minimal RMSE after transfer learning was applied, although slight differences were observed, as some instances were able to falsely improve the RMSE. This indicates that our proposed algorithm succeeded in improving the positioning accuracy by transferring knowledge from related source domains to the target domain. It can be observed that positive knowledge from the source domain was transferred to the prediction task. On the other hand, the Gaussian mixture model also achieved the lowest positioning accuracy.

Figure 14 shows the comparison of the accuracy of indoor positioning classifiers applied to dataset A with new feature spaces with 46 and 15 dimensions. It was found that the proposed algorithm has the lowest root mean square error, i.e., 32% fewer mislocalizations for mobile devices using the 41 new derived features. This indicates that the proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from related source domains to the target domain. It can be observed that not only a positive knowledge from the source domain transferred to the prediction task but also the computational cost and memory consumption were effectively improved by excluding irrelevant features from the analysis. However, the Gaussian Mixture Model (GMM) was found to have the lowest accuracy in indoor fingerprint positioning. It leads to mislocalization of mobile devices about 133% of the time and consumes huge computational costs. On the other side, the accuracy of indoor positioning classifiers applied to dataset A
with the reduced dimension of 15 features has also been evaluated. It was found that the proposed algorithm has the lowest root mean square error, i.e., predicts 25% fewer wrongly mobile devices. This indicates that the proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from related source domains to the target domain. It can be observed that not only positive knowledge from the source domain was transferred to the prediction task, but also the computational cost and memory usage were effectively improved by excluding irrelevant features from the analysis. One can observe the visible differences in positioning accuracy for both scenarios with different dimensions of 41 and 15 feature spaces. Those with 41 feature spaces seem to have better positioning accuracy; however, there are some features whose weight values are almost zero. Moreover, this confirms the claim that the measurements coming from the 26 access points that were excluded from the analysis due to the weighting values for the co-occurrence of these access points from both ranges of instances were insignificant, so the inclusion of these instances would falsely increase our positioning accuracy. On the other hand, the Gaussian mixture model (GMM) was found to have the lowest accuracy in locating indoor objects based on fingerprints. In about 137% of the cases, mobile devices are predicted incorrectly and high computational costs are incurred.

Figure 13. Comparisons of classifier’s RMSE for dataset A with 15 Wi-Fi APs.
Figure 14. Comparison of performance of the classifiers for dataset A with 41 and 15 APs.

Figure 15 illustrates the impact of different feature spaces used to train indoor positioning classifiers using dataset A with dimensions of 41 and 15, respectively. It was found that the proposed algorithm has the lowest root mean square error, i.e., 32% and 25% less error in positioning the mobile device for both scenarios. This indicates that the proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from the source domain to the target domain. It can be observed that not only positive knowledge from the source domain was transferred to the prediction task, but also the computational cost and memory consumption were effectively improved by excluding irrelevant features from the analysis. The difference in positioning accuracy between the two scenarios is about 7%. This difference led us to the important finding that all classifiers achieved lower RMSE both before and after transfer learning for the new feature spaces of dimension 41. However, the contributions of these features to target prediction were not as significant (since their co-occurrence weights of those 26 features were almost negligible) as for the feature spaces with dimension 15. This could show us researchers that a model with a lower RMSE does not mean that it is always the most consistent model, but that we need to examine the parameters for their negative effects on target prediction.

4.3.2. Dataset B

The original number of feature spaces was nine for dataset B. Figure 16 illustrates the distributions of the received signal strength measurements for a given Wi-Fi access point with their corresponding target values. It shows that the RSS measurements have a specific pattern for the different Wi-Fi access points and follow a different distribution of RSS values for the different labels of Wi-Fi AP one (AP_1) of dataset B. This shows that the signal variations of the Wi-Fi-received signal strengths were measured at a reference point and have different distribution values, including the standard deviation and the average value of the received signals. The unique pattern for the measurements collected from Wi-Fi AP (AP_1) also appears to have different values for the signal-to-location relationships of the different target points. Beyond that, there are no noticeable outlier values (as can be observed in Figure 15, except for the RSS values of $-90$, which demonstrate that the corresponding access points were not detected by any mobile devices).
Figure 15. Effect of different feature spaces used to train classifiers for dataset A.

Figure 16. Distributions of RSS for Wi-Fi-AP1 with their corresponding Targets for dataset B.

Figure 17 shows that the RSS measurements have a specific pattern for the different Wi-Fi access points and follow a different distribution of RSS values for the different labels of Wi-Fi AP_3 of dataset B. This shows that the signal variations of the Wi-Fi-received signal strengths were measured at a reference point and have different distribution values including the standard deviation and the average values of the received signals. The unique pattern for the measurements collected by the Wi-Fi AP_3 also appears to have different values for the relationships between the signal locations of the different target points. In addition, there are no noticeable outlier values (as can be observed in Figure 16, except for
the RSS values of \(-90\), which demonstrate that the corresponding APs were not detected by any mobile devices).

\[ \text{Figure 17. Distributions of RSS for Wi-Fi-AP3 with their corresponding targets for dataset B.} \]

Figure 18 shows that the RSS measurements for the different pair of access points have a certain pattern and follow the different distribution of RSS values for the different labels of an access point (dissimilarity). It is further described that the Wi-Fi signal strength received at a network point by multiple APs is assumed to be independent so that the Wi-Fi signals sent by different APs are transmitted independently and do not interfere with each other. Therefore, the distributions of RSS measurements for pairs of Wi-Fi access points demonstrate a different pattern of fingerprint for a mobile device at a given reference point even after performing transformation techniques. It can be observed that the RSS measurements for the different Wi-Fi access points of AP_3 and AP_7 have a specific pattern and follow the different distribution of RSS values for the different labels of Wi-Fi access points (dissimilarity). Thus, the relationship between signal and location has been characterized or studied and shows that the distributions of RSS values measured at a given network point are dynamic due to the inherent heterogeneity of the indoor environment. Thus, there is no guarantee that the signal variations at a particular location are represented by a single value.

Figure 19 describes whether the Wi-Fi signal strength received at a grid point from multiple APs can be assumed to be independent. Thus, the Wi-Fi signals transmitted by different APs are transmitted independently and do not interfere with each other. Therefore, the distributions of RSS measurements for pairs of Wi-Fi APs still follow a different pattern of fingerprint distribution for a mobile user at a given reference point after applying transformation techniques. It also shows that the RSS measurements for the different Wi-Fi access points of AP_8 and AP_9 have distinct patterns and follow the different distribution of RSS values for the different labels of a Wi-Fi access point. The “signal-to-location” relationship was established and further conveys that the distributions of RSS values measured at a given grid point are dynamic, and therefore there is no guarantee that the signal variations are represented by a single value at a given location.
of Wi-Fi access points (dissimilarity). Thus, the relationship between signal and location has been characterized or studied and shows that the distributions of RSS values measured at a given network point are dynamic due to the inherent heterogeneity of the indoor environment. Thus, there is no guarantee that the signal variations at a particular location are represented by a single value.

Figure 18. Distributions of RSS for Pair Wi-Fi AP_3 and AP_7 for dataset B.

Figure 19. Distributions of RSS for Pair Wi-Fi APs of AP_8 and AP_9 for dataset B.
Table 9 shows extensive experiments applied to real-world indoor positioning scenarios using dataset B. The original number of feature spaces was nine for dataset B. Nevertheless, experimental results confirmed that the proposed algorithm is the best algorithm with the minimum root mean square error for dataset B, regardless of which feature spaces were fully considered in the analysis. Since the number of feature spaces for dataset B is very small, we could not see or feel the actual effect of irrelevant features. However, based on our given scenarios, some redundant features need to be handled to minimize both computational cost and memory requirements. Experimental results have demonstrated that all classifiers achieved a minimum RMSE after applying transfer learning. This indicates that our proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from the source domain to the target domain. However, there is no clear evidence against the fact that negative knowledge was not transferred from the source domain to the target domain. In this case, the measurements that came from all Wi-Fi APs were not refined or reviewed to determine whether or not they should be used for location analysis, because they could negatively affect the location process. On the other hand, the Gaussian mixed model also achieved the lowest location accuracy and consumed higher computational costs.

Table 9. Models’ target prediction using the original feature spaces of 9 Wi-Fi access points based on OHetTLAL.

| Classifiers                        | RMSE (in Meters) |
|------------------------------------|------------------|
| Decision tree                      | 2.01             |
| K-Neighbour (KNN)                  | 1.89             |
| Support vector machine (SVC)       | 1.87             |
| Logistic Regression (LR)           | 1.96             |
| Random forest                      | 1.95             |
| Gaussian Mixture                   | 3.47             |
| Neural Network (MLP)               | 1.94             |
| The proposed algorithm (based TL)  | 1.25             |

Table 10 shows detailed experiments for real-world indoor positioning scenarios using dataset B. Here, we evaluated the positioning accuracy of the different predictive modeling techniques to estimate the target position of a mobile device based on the received signal strengths from different Wi-Fi APs. The proposed model was compared with seven machine learning algorithms. The data were preprocessed and irrelevant feature spaces were excluded from further analysis based on their weighted co-occurrence score to avoid negative knowledge transfer. The original number of feature spaces was nine for dataset B before feature transformation was being carried out. However, based on the co-occurrence weight scores and exclusion criteria, three feature spaces were deemed irrelevant for analysis because the co-occurrence weight scores were zero in each of the specified base stations. In addition, the Wi-Fi-received signal strength measurements that achieved the lowest co-occurrence weight scores for the offline and online source instances were excluded from the prediction task because they would have no effect on the position estimation. Accordingly, as shown in Table 9 for dataset B, only six Wi-Fi APs are used as new feature spaces to train the classifiers. In this case, the measurements that came from the remaining Wi-Fi APs were not included in the analysis because they could have a negative impact on the positioning process, since the co-occurrence weight scores of these APs from both domains of the instances were insignificant, and therefore including them would falsely inflate our positioning accuracy. The experimental results confirm that the proposed algorithm is also the best algorithm for dataset B with the lowest root mean square error. This indicates that the proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from related source domains to the target domain. It can be observed that not only a positive knowledge was transferred from the source domain to the prediction task, but also the computational cost and memory usage were
effectively improved by excluding irrelevant features from the analysis. On the other hand, the Gaussian mixture model has achieved the lowest positioning accuracy and consumes a high computational cost.

Table 10. Models’ target prediction using the new feature spaces of 6 Wi-Fi access points based on OHetTLAL.

| Classifiers                  | RMSE (in Meters) |
|------------------------------|------------------|
| Decision tree                | 1.95             |
| K -Neighbour (KNN)          | 1.93             |
| Support vector machine (SVC)| 1.83             |
| Logistic Regression (LR)    | 2.24             |
| Random forest                | 1.92             |
| Gaussian Mixture             | 3.58             |
| Neural Network (MLP)        | 1.91             |
| The proposed algorithm (based TL) | 1.42         |

Figure 20 shows that the proposed algorithm is the best algorithm with the lowest mean square error for dataset B, with nine Wi-Fi APs. On the other hand, the Gaussian mixture model has also achieved the lowest positioning accuracy. All classifiers have achieved minimum RMSE after applying transfer learning. This indicates that our proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from the source domain to the target domain. However, there is no strong evidence against the fact that a negative knowledge was not transferred from the source domain to the target domain. In this case, the measurements coming from all access points were not refined or analyzed to determine whether or not they should be used for location analysis because they could negatively impact the positioning process.

Figure 21 shows that the proposed algorithm was the best algorithm with the lowest mean square error for dataset B with six feature spaces. In this case, the measurements coming from the three Wi-Fi APs were not included in the analysis because they could negatively affect the prediction task. This is because the coincidence weights of these Wi-Fi access points from both domains are close to zero, so including them would affect the positioning performance. Similarly, all classifiers achieved minimum RMSE after applying transfer learning. This indicates that our proposed algorithm has succeeded in improving
the positioning performance by transferring knowledge from the corresponding source domain to the target domain. It can be observed that positive knowledge from the source domain was transferred to the prediction task. On the other hand, the Gaussian mixture model also achieved the lowest positioning accuracy.

Figure 21. Comparisons of classifier’s RMSE (in meter) for dataset B with 6 Wi-Fi APs.

Figure 22 describes the comparison of the accuracy of indoor positioning classifiers applied to dataset B with nine feature spaces. It was found that the proposed algorithm has the lowest root mean square error, i.e., 33% less mis-localized mobile objects. This indicates that the proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from related source spaces to the target area. However, there is no clear evidence against the fact that negative knowledge was not transferred from the source to the target domain. In this case, the measurements that came from all Wi-Fi access points were not refined or reviewed to determine whether or not they should be used for location analysis, because they could negatively affect the positioning process. The Gaussian mixture model (GMM), however, was found to have the lowest accuracy for indoor fingerprint-based location. In about 86% of the cases, mobile devices are incorrectly localized and high computational costs are incurred. Moreover, the comparison of the accuracy of indoor positioning classifiers applied to dataset B have also been evaluated with six feature spaces. It was found that the proposed algorithm has the lowest mean square error, i.e., 22% fewer mis-localizations for mobile objects. This indicates that the proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from related source domains to the target domain. It can be observed that not only a positive knowledge was transferred from the source domain to the prediction task, but also the computational cost and memory usage were effectively improved by excluding irrelevant features from the analysis. Similarly, the Gaussian mixture model (GMM) was found to have the lowest accuracy for indoor fingerprint-based location. In about 96% of the cases, it results in the incorrect positioning of mobile devices and high computational complexity.
leads us to an important finding, namely that all classifiers achieved lower root mean square error, i.e., 1.25 m and 1.42 m less mis-localizations of mobile devices for both scenarios. This indicates that the proposed algorithm has succeeded in improving the location accuracy by transferring knowledge from the source domain to the target domain, where the difference in location accuracy between the two scenarios is about 11%. This difference leads us to an important finding, namely that all classifiers achieved lower RMSE both before and after transfer learning for the new feature spaces of dimension 9. However, there is no strong evidence that negative knowledge was not transferred from the source to the target domain. In this case, especially for the scenario of dataset A with nine feature spaces, the measurements coming from all Wi-Fi access points were not refined or reviewed to determine whether or not they should be used for location analysis because they could negatively affect the positioning process. Moreover, the contributions of these features to target detection were not as significant (since their co-occurrence weight scores of those three features were almost negligible) as for the feature spaces with a dimension of 6. This could justify that a model with a lower RMSE does not mean that it is always the most consistent model, but that we need to examine the parameters for their negative effects on target prediction.

**Figure 22.** Comparison of performance of the classifiers for dataset B with 9 and 6 APs.

**Figure 23.** Effect of different feature spaces used to train classifiers for dataset B.
Figure 24a,b and Figure 25a,b illustrate the use of box and whisker (commonly known as a box plot) to compare the distributions of the RSS measurements received from an access point in a reference point of both training and testing instances, respectively. For both the training and testing instances, we have considered the measurements received from Wi-Fi AP_1 and Wi-Fi AP_48 to compare the nature of the signal distributions in terms of the box-plot parameters. One can observe that the signal distributions of reference point 1 for Wi-Fi AP_1 seem to have outlier values (both types of outliers: low and high), because the RSS values lies out of the whisker box such that the RSS measurements either fall more than 1.5 times the interquartile range above the third quartile or fall more than 1.5 times the interquartile range below the first quartile. Inter quartile range is the difference between the third and first quartile, which used to give the dispersion of a distribution. On the other side, Targets of 185, 191 and 210 have a lower outlier. One can also see that the respective medians of each box plot of a target lies outside of any box of a comparison boxplot of any target. Thus, the unique pattern for the measurements collected by the Wi-Fi APs seems to have different median values for the ‘signal-to-location’ relationships of the different target points. In other words, the received signals for a reference point from multiple APs are likely to be independent and it is what is really required in practice for reducing extra Wi-Fi APs’ deployment. Similarly, one can compare the distribution nature in terms of skewness (positive and negative), symmetry, potential outliers, and dispersion of measurements. For instance, the measurements for targets 5 and 10 of Wi-Fi AP_4 of both training and testing instances revealed that the destruction is highly dispersed and both distributions are negatively skewed. Moreover, the measurements’ distribution received from Wi-Fi access points in both phases of instances are highly dispersed as the whisker box plot is longer. This also justifies the fact that signals were not received from this access point for their corresponding targets with the exception of a few measurements, while this automatically inflates the deviation and degrades the positioning performance.

Figure 24. Distribution analysis of Wi-Fi RSS training measurements of different APs based on box and whisker plot. The small circles found above and below the whisker box plot represents the outliers of the training measurements received from the Wi-Fi APs of (a) AP_1 and (b) AP_48.

4.4. Analysis of Computational Complexity

Figure 26 presents the computational complexity of the proposed algorithm for IPS-based RSS fingerprinting, which fuses source and target domain knowledge to enhance positioning performance in the target domain. The algorithms presented in this paper were run on a laptop with an AMD Ryzen 3 3200U CPU running at 2.60 GHz and 16 GB of RAM. Two situations of complexity—time and space—are the main determinants of an algorithm’s complexity. To assess the time complexity of the algorithms, we employed both the functional analysis of Big O and the typical execution testing time. The overall worst-case computational cost analysis is given by the Big O functional analysis of time complexity,
and all of the algorithms we used have an average of worst-case time complexity of $O(n^2)$. However, compared to the other algorithms we used, the proposed algorithm’s average elapsed execution testing time is substantially longer. Additionally, employing the two real-world dataset scenarios, Figure 26 shows the comparative study of computational testing time for FBIPS. Thus, the computational testing time for the proposed algorithm is higher, followed by the GMM, MLP, and the SVC.

![Box and whisker plot](image)

**Figure 25.** Distribution analysis of Wi-Fi RSS testing measurements of different APs based on box and whisker plot. The small circles found above and below the whisker box plot represents the outliers of the testing measurements received from the Wi-Fi APs of (a) AP_1 and (b) AP_48.

![Graph](image)

**Figure 26.** Comparative analysis of computational testing time of the algorithms used for FBIPS.

5. Conclusions

In this paper, we propose OHetTLAL, a novel technique for indoor fingerprint-based positioning problems that can improve learning performance in the target domain through knowledge transfer from the source domain. To this end, we derived new feature spaces (on which the model is trained) based on the co-occurrence of RSS measurements from mobile devices in the two domains to capitalize on knowledge that could improve target location prediction. To derive the new feature spaces, the co-occurrence measure of the features ($C_{int}^P$) was computed. The higher the value of co-occurrence between the two
domains, the more detected Wi-Fi APs are shared by these two domains and the more likely these source domains are related to the target domain, which positively affects the prediction of the target. The proposed objective function was minimized across the new feature spaces so that the positive knowledge of the related source domain received a higher weight and is transferred to the target domain for predicting a new mobile device. Experimental results demonstrate that the proposed algorithm was the best algorithm for fingerprint-based indoor positioning with the minimum root mean square error in all cases of the two real office scenarios. On the other hand, the Gaussian mixture model (GMM), had the highest mis-localization error rate. Our research shows that in OHetTL, the construction of new feature spaces is critical for extracting relevant information for effective target location prediction. It should be noted that these experimental results may differ in other test phase environments.

The effects of different dimensions of feature spaces used for training indoor positioning classifiers have also been evaluated using two real life datasets. Results demonstrate that the impact of different feature spaces used to train indoor positioning classifiers using dataset A with dimensions of 41 and 15, respectively. It was found that the proposed algorithm has the lowest root mean square error, i.e., 32% and 25% less error in positioning the mobile device for both scenarios. Similarly, the effects of different dimensions of feature spaces are used for training indoor positioning classifiers, using dataset B with a dimension of 9 and 6, respectively. It was found that the proposed algorithm has the lowest root mean square error, i.e., 33% and 22% less mis-localizations of mobile devices for both scenarios. This indicates that the proposed algorithm has succeeded in improving the positioning accuracy by transferring knowledge from the source domain to the target domain for both datasets after new feature spaces are derived based on the weight co-occurrence of both instances. It can be observed that not only positive knowledge from the source domain was transferred to the prediction task, but also the computational cost and memory consumption were effectively improved by excluding irrelevant features from the analysis. The difference in positioning accuracy obtained due to the feature spaces between the two scenarios is about 7% and 11% for both datasets of A and B, respectively. This difference leads us to an important finding, namely that all classifiers achieved lower RMSE both before and after transfer learning over the new feature spaces of dimension 9. However, there is no strong evidence that negative knowledge was not transferred from the source to the target domain. In this case, especially for the scenario of dataset B with nine feature spaces, the measurements coming from all Wi-Fi access points were not refined or scrutinized to determine whether or not they should be used for location analysis because they could negatively affect the positioning process. However, the contributions of these features to target detection were not as significant (since their co-occurrence weight scores of those three features were almost negligible) as for the feature spaces with a dimension of 6. This could justify that a model with a lower RMSE does not mean that it is always the most consistent model, but that we need to examine the parameters for their negative effects on target prediction. Our results demonstrate, however, that the proposed algorithm is more resilient to fluctuating environments and mitigates the problem of overfitting the model. Furthermore, the OHetTLAL-based fusion of both knowledge domains could significantly improve the positioning performance without any extra cost associated to offline calibrations effort. It is not necessarily the positioning performance to improve as the number of feature spaces increases, instead, how the features are important in capturing the maximum variation of the model is the main determinant.

In summary, the positioning performance, as per our results, revealed that it is highly important to extract the most significant predictors that could capture the maximum variance of the model using some metrics in order to evaluate the significance of each Wi-Fi AP deployment. Thus, we discovered that not all available feature spaces need be considered for positioning estimation for at least four reasons: (a) some features may be irrelevant and result in fingerprint duplication or pattern fingerprint mismatch; (b) large feature spaces necessitate a massive deployment of Wi-Fi access points, which is prohibitively expensive
from a cost standpoint; (c) from a technical standpoint, computational cost and memory usage are quite expensive when dealing with higher dimensions of features, and (d) model over-fitting is also a serious issue for higher dimensions of feature spaces.

**Author Contributions:** Conceptualization, H.T.G., X.G., K.Z., L.L. and Y.Z.; Data curation, L.L. and Y.Z.; Formal analysis, H.T.G., X.G. and K.Z.; Funding acquisition, X.G.; Investigation, H.T.G., X.G. and K.Z.; Methodology, H.T.G., X.G. and K.Z.; Project administration, X.G. and K.Z.; Resources, L.L. and Y.Z.; Supervision, X.G.; Visualization, K.Z. and Y.Z.; Writing—original draft, H.T.G.; Writing—review & editing, X.G., K.Z., L.L. and Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the National Natural Science Foundation of China under Grant (No. 62171086) and in part by the Municipal Government of Quzhou under Grant (No. 2021D002).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The dataset used for this study are available upon request to the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Guo, X.; Chu, L.; Sun, X. Accurate localization of multiple sources using semidefinite programming based on incomplete range matrix. *IEEE Sens. J.* 2016, 16, 5319–5324. [CrossRef]

2. Moghtadaiee, V.; Dempster, A.G. WiFi fingerprinting signal strength error modeling for short distances. In Proceedings of the 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sydney, Australia, 13–15 November 2012; pp. 1–6.

3. Patwari, N. *Location Estimation in Sensor Networks*; University of Michigan: Ann Arbor, MI, USA, 2005.

4. Bahl, P.; Padmanabhan, V.N. RADAR: An in-building RF-based user location and tracking system. In Proceedings of the IEEE INFOCOM 2000, Conference on Computer Communications, Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064), Tel Aviv, Israel, 26–30 March 2000; Volume 2, pp. 775–784.

5. Bajaj, R.; Ranaweera, S.L.; Agrawal, D.P. GPS: Location-tracking technology. *Computer* 2002, 35, 92–94. [CrossRef]

6. Guo, X.; Ansari, N. Localization by fusing a group of fingerprints via multiple antennas in indoor environment. *IEEE Trans. Veh. Technol.* 2017, 66, 9904–9915. [CrossRef]

7. Sun, X.; Ansari, N. EdgeIoT: Mobile edge computing for the Internet of Things. *IEEE Commun. Mag.* 2016, 54, 22–29. [CrossRef]

8. Guo, X.; Ansari, N.; Li, L.; Duan, L. A hybrid positioning system for location-based services: Design and implementation. *IEEE Commun. Mag.* 2020, 58, 90–96. [CrossRef]

9. Rodriguez, M.D.; Favela, J.; Martínez, E.A.; Muñoz, M.A. Location-aware access to hospital information and services. *IEEE Trans. Inf. Technol. Biomed.* 2004, 8, 448–455. [CrossRef]

10. Harle, R. A survey of indoor inertial positioning systems for pedestrians. *IEEE Commun. Surv. Tutor.* 2013, 15, 1281–1293. [CrossRef]

11. Patterson, C.A.; Munzt, R.R.; Pancake, C.M. Challenges in location-aware computing. *IEEE Pervasive Comput.* 2003, 2, 80–89. [CrossRef]

12. Tili, F.; Hamdi, N.; Belghith, A. Accurate 3D localization scheme based on active RFID tags for indoor environment. In Proceedings of the 2012 IEEE International Conference on RFID-Technologies and Applications (RFID-TA), Nice, France, 5–7 November 2012; pp. 378–382.

13. Luo, Q.; Yang, J.; Sui, Z.; Ai, Q. Bluetooth indoor positioning combination algorithm based on fingerprint matching technology. In Proceedings of the 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Fuzhou, China, 24–26 September 2021; pp. 35–40.

14. Poulose, A.; Han, D.S. UWB indoor localization using deep learning LSTM networks. *Appl. Sci.* 2020, 10, 6290. [CrossRef]

15. Guo, X.S.; Chu, L.; Li, B.; Xu, B.; Wan, Q.; Shen, Y. A robust vector matching localization approach based on multiple channels SSD fingerprinting of Zigbee networks. *Prog. Electromagn. Res.* 2014, 144, 133–140.

16. Poulose, A.; Eyobu, O.S.; Han, D.S. An indoor position-estimation algorithm using smartphone IMU sensor data. *IEEE Access* 2019, 7, 11165–11177. [CrossRef]

17. Guo, X.; Shao, S.; Ansari, N.; Khreishah, A. Indoor localization using visible light via fusion of multiple classifiers. *IEEE Photonics J.* 2017, 9, 1–16. [CrossRef]

18. AlNuaimi, K.; Kamel, H. A survey of indoor positioning systems and algorithms. In Proceedings of the 2011 International Conference on Innovations in Information Technology (IIT), Abu Dhabi, United Arab Emirates, 25–27 April 2011; pp. 185–190.
19. Collin, J.; Mezentsev, O.; Lachapelle, G. Indoor positioning system using accelerometers and high accuracy heading sensors. In Proceedings of the ION GPS/GNSS 2003 Conference, Portland, OR, USA, 9–12 September 2003; pp. 9–12.

20. Alarifi, A.; Al-Salman, A.; Alsaleh, M.; Alnafessah, A.; Al-Hadhrami, S.; Al-Ammar, M.A.; Al-Khalifa, H.S. Ultra wideband indoor positioning technologies: Analysis and recent advances. Sensors 2016, 16, 707. [CrossRef] [PubMed]

21. Filippoupolitis, A.; Oliff, W.; Loukas, G. Bluetooth low energy based occupancy detection for emergency management. In Proceedings of the 2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberpace and Security (UCC-CSS), New York, NY, USA, 14–16 December 2016; pp. 31–38.

22. Tekler, Z.D.; Low, R.; Yuen, C.; Blessing, L. Plug-Mate: An IoT-based occupancy-driven plug load management system in smart buildings. Build. Environ. 2022, 223, 109472.

23. Balaji, B.; Xu, J.; Nwokoro, A.; Gupta, R.; Agarwal, Y. Sentinel: Occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings. In Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, New York, NY, USA, 10–13 November 2013; pp. 1–14.

24. Wang, S.S. A BLE-based pedestrian navigation system for car searching in indoor parking garages. Sensors 2018, 18, 1442. [CrossRef]

25. Filippoupolitis, A.; Oliff, W.; Takand, B.; Loukas, G. Location-enhanced activity recognition in indoor environments using off the shelf smart watch technology and BLE beacons. Sensors 2017, 17, 1230. [CrossRef]

26. Tekler, Z.D.; Low, R.; Blessing, L. An alternative approach to monitor occupancy using bluetooth low energy technology in an office environment. In Journal of Physics: Conference Series; IOP Publishing: Bristol, UK, 2019; Volume 1343, p. 012116.

27. Ng, P.C.; Spachos, P.; Plataniotis, K.N. COVID-19 and your smartphone: BLE-based smart contact tracing. IEEE Syst. J. 2021, 15, 5367–5378. [CrossRef] [PubMed]

28. Chang, K.H. Bluetooth: A viable solution for IoT? [Industry Perspectives]. IEEE Wirel. Commun. 2014, 21, 6–7. [CrossRef]

29. Guo, X.; Li, L.; Xu, F.; Ansari, N. Expectation maximization indoor localization utilizing supporting set for Internet of Things. IEEE Internet Things J. 2018, 6, 2573–2582. [CrossRef]

30. Alsindi, N.A.; Alavi, B.; Pahlavan, K. Measurement and modeling of ultrawideband TOA-based ranging in indoor multipath environments. IEEE Trans. Veh. Technol. 2008, 58, 1046–1058. [CrossRef]

31. Ma, Y.; Wang, B.; Pei, S.; Zhang, Y.; Zhang, S.; Yu, J. An indoor localization method based on AOA and PDOA using virtual stations in multipath and NLOS environments for passive UHF RFID. IEEE Access 2018, 6, 31772–31782. [CrossRef]

32. Huang, B.; Xie, L.; Yang, Z. TDOA-based source localization with distance-dependent noises. IEEE Trans. Wirel. Commun. 2014, 14, 468–480. [CrossRef]

33. Liu, H.; Darabi, H.; Banerjee, P.; Liu, J. Survey of wireless indoor positioning techniques and systems. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 2007, 37, 1067–1080. [CrossRef]

34. Chen, L.; Yang, K.; Wang, X. Robust cooperative Wi-Fi fingerprint-based indoor localization. IEEE Internet Things J. 2016, 3, 1406–1417. [CrossRef]

35. Sun, J.; He, L.; Gu, Y.; Jiang, W.; Kushwaha, G.; Vinip, A.; Cheng, L.; Cong, L.; Zhu, T. Low-overhead WiFi fingerprinting. IEEE Trans. Comput. 2017, 590–603. [CrossRef]

36. Guo, X.; Li, L.; Ansari, N.; Liao, B. Knowledge aided adaptive localization via global fusion profile. IEEE Internet Things J. 2017, 5, 1081–1089. [CrossRef]

37. He, S.; Chan, S.H.G. Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. IEEE Commun. Surv. Tutor. 2015, 18, 466–490. [CrossRef]

38. Roos, T.; Myllymäki, P.; Pirri, H.; Misikangas, P.; Sievänen, J. A probabilistic approach to WLAN user location estimation. Int. J. Wirel. Inf. Netw. 2002, 9, 155–164. [CrossRef]

39. Fang, S.H.; Lin, T.N.; Lee, K.C. A novel algorithm for multipath fingerprinting in indoor WLAN environments. IEEE Trans. Wirel. Commun. 2008, 7, 3579–3588. [CrossRef]

40. Wu, Z.L.; Li, C.H.; Ng, J.K.Y.; Leung, K.R. Location estimation via support vector regression. IEEE Trans. Mob. Comput. 2007, 6, 311–321. [CrossRef]

41. Poulose, A.; Han, D.S. Performance analysis of fingerprint matching algorithms for indoor localization. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIC), Fukuoka, Japan, 19–21 February 2020; pp. 661–665.

42. Kotaru, M.; Joshi, K.; Bharadia, D.; Katti, S. Spotfi: Decimeter level localization using wifi. In Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication, London, UK, 17–21 August 2015; pp. 269–282.

43. Guo, X.; Huang, Y.; Li, B.; Chu, L. DOA estimation of mixed circular and non-circular signals using uniform circular array. In Proceedings of the 2014 7th International Conference on Image and Signal Processing, Dalian, China, 20–23 October 2014; pp. 1043–1047.

44. Otim, T.; Bahillo, A.; Diez, L.E.; Lopez-Iturri, P.; Falcone, F. FDTD and empirical exploration of human body and UWB radiation interaction on TOF ranging. IEEE Antennas Wirel. Propag. Lett. 2019, 18, 1119–1123. [CrossRef]

45. Guo, X.; Ansari, N.; Li, L.; Li, H. Indoor localization by fusing a group of fingerprints based on random forests. IEEE Internet Things J. 2018, 5, 4686–4698. [CrossRef]

46. Poulose, A.; Kim, J.; Han, D.S. A sensor fusion framework for indoor localization using smartphone sensors and Wi-Fi RSSI measurements. Appl. Sci. 2019, 9, 4379. [CrossRef]
47. Poulouse, A.; Han, D.S. Hybrid indoor localization using IMU sensors and smartphone camera. *Sensors* 2019, 19, 5084. [CrossRef]  
48. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 2009, 22, 1345–1359. [CrossRef]  
49. Sun, Z.; Chen, Y.; Qi, J.; Liu, J. Adaptive localization through transfer learning in indoor Wi-Fi environment. In *Proceedings of the 2008 Seventh International Conference on Machine Learning and Applications*, San Diego, CA, USA, 11–13 December 2008; pp. 331–336.  
50. Guo, X.; Ansari, N.; Hu, F.; Shao, Y.; Elikplim, N.R.; Li, L. A survey on fusion-based indoor positioning. *IEEE Commun. Surv. Tutor.* 2019, 22, 566–594. [CrossRef]  
51. Tekler, Z.D.; Low, R.; Gunay, B.; Andersen, R.K.; Blessing, L. A scalable Bluetooth Low Energy approach to identify occupancy patterns and profiles in office spaces. *Build. Environ.* 2020, 171, 106681. [CrossRef]  
52. Baronti, P.; Barsocchi, P.; Chessa, S.; Mavilia, F.; Palumbo, F. Indoor Bluetooth low energy dataset for localization, tracking, occupancy, and social interaction. *Sensors* 2018, 18, 4462. [CrossRef]  
53. Zhang, S.; Guo, J.; Wang, W.; Hu, J. Indoor 2.5 D positioning of WiFi based on SVM. In *Proceedings of the 2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS)*, Wuhan, China, 22–23 March 2018; pp. 1–7.  
54. Li, W.; Wei, D.; Yuan, H.; Ouyang, G. A novel method of WiFi fingerprint positioning using spatial multi-points matching. In *Proceedings of the 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Madrid, Spain, 4–7 October 2016; pp. 1–8.  
55. Yu, H.K.; Oh, S.H.; Kim, J.G. AI based location tracking in WiFi indoor positioning application. In *Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIC)*, Fukuoka, Japan, 27 February 2020; pp. 199–202.  
56. Alhammadi, A.; Alraih, S.; Hashim, F.; Rasid, M.F.A. Robust 3D indoor positioning system based on radio map using Bayesian network. In *Proceedings of the 2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, Limerick, Ireland, 15–18 April 2019; pp. 107–110.  
57. Wu, P.; Dietterich, T.G. Improving SVM accuracy by training on auxiliary data sources. In *Proceedings of the Twenty-First International Conference on Machine Learning*, Banff, AB, Canada, 4–8 July 2004; p. 110.  
58. Liu, X.; Zhan, Y.; Cen, J. An energy-efficient crowd-sourcing-based indoor automatic localization system. *IEEE Sens. J.* 2018, 18, 6009–6022. [CrossRef]  
59. Ferris, B.; Fox, D.; Lawrence, N.D. Wifi-slam using gaussian process latent variable models. In *Proceedings of the IJCAI*, Hyderabad, India, 6–12 January 2007; Volume 7, pp. 2480–2485.  
60. Alraih, S.; Alhammadi, A.; Shaye, I.; Al-Samman, A.M. Improving accuracy in indoor localization system using fingerprinting technique. In *Proceedings of the 2017 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Republic of Korea, 18–20 October 2017; pp. 274–277.  
61. Brunato, M.; Kiss Kallo, C. Transparent Location Fingerprinting for Wireless Services. 2002. Available online: [http://eprints.biblio.unitn.it/227/](http://eprints.biblio.unitn.it/227/) (accessed on 27 July 2021).  
62. King, T.; Kopf, S.; Haenselmann, T.; Luberger, C.; Effelsberg, W. Compass: A probabilistic indoor positioning system based on 802.11 and digital compasses. In *Proceedings of the 1st International Workshop on Wireless Network Testbeds*, Experimental Evaluation & Characterization, Los Angeles, CA, USA, 29 September 2006; pp. 34–40.  
63. Zou, D.; Meng, W.; Han, S.; Gong, Z.; Yu, B. User aided self-growing approach on radio map construction for WLAN based localization. In *Proceedings of the 26th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2013)*, Nashville, TN, USA, 16–20 September 2013; pp. 991–997.  
64. Laariedh, M.; Ugben, B.; Stephan, J.; Corre, Y.; Lostanlen, Y.; Raspopoulos, M.; stavrou, S. Ray tracing-based radio propagation modeling for indoor localization purposes. In *Proceedings of the 2012 IEEE*, Barcelona, Spain, 17–19 September 2012.  
65. Faragher, R.M.; Harle, R.K. SmartSLAM-an efficient smartphone indoor positioning system exploiting machine learning and opportunistic sensing. In *Proceedings of the 26th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2013)*, Nashville, TN, USA, 16–20 September 2013; pp. 1006–1019.  
66. Zhu, X.; Ghahramani, Z. Learning from Labeled and Unlabeled Data with Label Propagation. 2002. Available online: [https://www.researchgate.net/publication/2475534_Learning_from_Labeled_and_Unlabeled_Data_with_Label_Propagation](https://www.researchgate.net/publication/2475534_Learning_from_Labeled_and_Unlabeled_Data_with_Label_Propagation) (accessed on 21 July 2021).  
67. Eaton, E. Selective transfer between learning tasks using task-based boosting. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, 11 August 2011.  
68. Dai, W.; Xue, G.R.; Yang, Q.; Yu, Y. Transferring naive bayes classifiers for text classification. In *Proceedings of the AAAI*, Vancouver, BC, Canada, July 22–26 2007; Volume 7, pp. 540–545.  
69. Ding, Z.; Shao, M.; Fu, Y. Deep low-rank coding for transfer learning. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, Buenos Aires, Argentina, 25–31 June 2015.  
70. Shao, M.; Kit, D.; Fu, Y. Generalized transfer subspace learning through low-rank constraint. *Int. J. Comput. Vis.* 2014, 109, 74–93. [CrossRef]  
71. Pan, W.; Yang, Q. Transfer learning in heterogeneous collaborative filtering domains. *Artif. Intell.* 2013, 197, 39–55. [CrossRef]  
72. Zhou, J.T.; Tsang, I.W.; Pan, S.J.; Tan, M. Heterogeneous domain adaptation for multiple classes. In *Proceedings of the Artificial Intelligence and Statistics*, New York, NY, USA, 24 June 2014; pp. 1095–1103.
73. Dai, W.; Chen, Y.; Xue, G.R.; Yang, Q.; Yu, Y. Translated learning: Transfer learning across different feature spaces. *Adv. Neural Inf. Process. Syst.* **2008**, *21*, 353–360.

74. Wang, G.; Hoiem, D.; Forsyth, D. Building text features for object image classification. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 1367–1374.

75. Ng, M.K.; Wu, Q.; Ye, Y. Co-transfer learning via joint transition probability graph based method. In Proceedings of the 1st International Workshop on Cross Domain Knowledge Discovery in Web and Social Network Mining, New York, NY, USA, 12 August 2012; pp. 1–9.

76. Wu, Q.; Ng, M.K.; Ye, Y. Cotransfer learning using coupled Markov chains with restart. *IEEE Intell. Syst.* **2013**, *29*, 26–33. [CrossRef]

77. Yang, L.; Jing, L.; Yu, J.; Ng, M.K. Learning transferred weights from co-occurrence data for heterogeneous transfer learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *29*, 26–33. [CrossRef]

78. Xiang, E.W.; Pan, S.J.; Pan, W.; Su, J.; Yang, Q. Source-selection-free transfer learning. In Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, Barcelona, Spain, 16–22 June 2011.

79. Shalev-Shwartz, S.; Singer, Y. Online Learning: Theory, Algorithms, and Applications. 2007. Available online: https://www.semanticscholar.org/paper/Online-Learning%3A-Theory%2C-Algorithms%2C-and-Shalev-Shwartz/fb54e7be71c825295dd8489c0b47ed43932cb817 (accessed on 12 July 2021).

80. Ge, L.; Gao, J.; Zhang, A. OMS-TL: A framework of online multiple source transfer learning. In Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, New York, NY, USA, 27 October 2013; pp. 2423–2428.

81. Wang, B.; Pineau, J. Online boosting algorithms for anytime transfer and multitask learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 7–11 February 2015; Volume 29.

82. Yin, J.; Yang, Q.; Ni, L.M. Learning adaptive temporal radio maps for signal-strength-based location estimation. *IEEE Trans. Mob. Comput.* **2008**, *7*, 869–883. [CrossRef]

83. Zheng, V.W.; Xiang, E.W.; Yang, Q.; Shen, D. Transferring Localization Models over Time. In Proceedings of the AAAI, Chicago, IL, USA, 12–17 July 2008; Volume 2008, pp. 13–17 July 2008; Volume 2008, pp. 1421–1426.

84. Zheng, V.W.; Pan, S.J.; Yang, Q.; Pan, J.J. Transferring Multi-device Localization Models using Latent Multi-task Learning. In Proceedings of the AAAI, Chicago, IL, USA, July 2013; Volume 2013, pp. 1427–1432.

85. Azar, M.G.; Lazaric, A.; Brunskill, E. Sequential transfer in multi-armed bandit with finite set of models. *arXiv* 2013, arXiv:1307.6887.

86. Sindhwani, V.; Rosenberg, D.S. An RKHS for multi-view learning and manifold co-regularization. In Proceedings of the 25th International Conference on Machine Learning, New York, NY, USA, 28–30 July 2008; pp. 976–983.

87. Tang, J.; Alelyani, S.; Liu, H. Feature selection for classification: A review. *Data Classif. Algorithms Appl.* **2014**, *37*, 28.

88. Zhao, P.; HOI, S.C. OTL: A framework of Online Transfer Learning. In Proceedings of the 27th International Conference on Machine Learning (ICML 2010), Haifa, Israel, 21–24 June 2010; pp. 1–219.

89. Zhao, P.; Hoi, S.C.; Wang, J.; Li, B. Online transfer learning. *Artif. Intell.* **2014**, *216*, 76–102. [CrossRef]

90. Smola, A.; Gretton, A.; Song, L.; Schölkopf, B. A Hilbert space embedding for distributions. In Proceedings of the International Conference on Algorithmic Learning Theory, Banff, AB, Canada, 29–31 October 2007; Springer: Berlin/Heidelberg, Germany; pp. 13–31.

91. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* **2010**, *22*, 199–210. [CrossRef]

92. Chérie-Abdellatif, B.E.; Alquier, P. MMD-Bayes: Robust Bayesian estimation via maximum mean discrepancy. In Proceedings of the Symposium on Advances in Approximate Bayesian Inference, Vancouver, BC, Canada, 1 February 2020; pp. 1–21.