



Jian Huang<sup>1</sup>, Haonan Si<sup>1,2</sup>, Xiansheng Guo<sup>1,2,\*</sup> and Ke Zhong<sup>1</sup>

- <sup>1</sup> Department of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
- <sup>2</sup> Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou 324000, China
- \* Correspondence: xsguo@uestc.edu.cn

Abstract: Distribution discrepancy is an intrinsic challenge in existing fingerprint-based indoor positioning system(s) (FIPS) due to real-time environmental variations; thus, the positioning model needs to be reconstructed frequently based on newly collected training data. However, it is expensive or impossible to collect adequate training samples to reconstruct the fingerprint database. Fortunately, transfer learning has proven to be an effective solution to mitigate the distribution discrepancy, enabling us to update the positioning model using newly collected training data in real time. However, in practical applications, traditional transfer learning algorithms no longer act well to feature space heterogeneity caused by different types or holding postures of fingerprint collection devices (such as smartphones). Moreover, current heterogeneous transfer methods typically require enough accurately labeled samples in the target domain, which is practically expensive and even unavailable. Aiming to solve these problems, a heterogeneous transfer learning framework based on co-occurrence data (HTL-CD) is proposed for FIPS, which can realize higher positioning accuracy and robustness against environmental changes without reconstructing the fingerprint database repeatedly. Specifically, the source domain samples are mapped into the feature space in the target domain, then the marginal and conditional distributions of the source and target samples are aligned in order to minimize the distribution divergence caused by collection device heterogeneity and environmental changes. Moreover, the utilized co-occurrence fingerprint data enables us to calculate correlation coefficients between heterogeneous samples without accurately labeled target samples. Furthermore, by resorting to the adopted correlation restriction mechanism, more valuable knowledge will be transferred to the target domain if the source samples are related to the target ones, which remarkably relieves the "negative transfer" issue. Real-world experimental performance implies that, even without accurately labeled samples in the target domain, the proposed HTL-CD can obtain at least 17.15% smaller average localization errors (ALEs) than existing transfer learning-based positioning methods, which further validates the effectiveness and superiority of our algorithm.

Keywords: indoor positioning; heterogeneous transfer learning; co-occurrence data

## 1. Introduction

An indoor positioning system is a vital technique as it enables a series of position-based Internet of things (IoT) applications, including smart cities [1,2], healthcare [3], emergency rescue [4,5], occupancy prediction [6,7], and intelligent buildings [8–10]. Generally, accurate and real-time position data are required in the aforementioned GNSS-denied scenarios, which are usually hindered by complex indoor structures, real-time environmental changes, multipath interferences, and so on. Currently, there are four fundamental indoor positioning solutions for these issues, i.e., geometric-based methods [11,12], vision-based [13] methods, pedestrian dead reckoning (PDR) [14], and fingerprint-based methods [15,16]. Among them, the fingerprint-based method receives the most attention due to its flexible



Citation: Huang, J.; Si, H.; Guo, X.; Zhong, K. Co-Occurrence Fingerprint Data-Based Heterogeneous Transfer Learning Framework for Indoor Positioning. *Sensors* **2022**, *22*, 9127. https://doi.org/10.3390/s22239127

Academic Editor: David Plets

Received: 18 October 2022 Accepted: 21 November 2022 Published: 24 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). implementations and easy availability of fingerprint samples using smart devices. In typical fingerprint-based indoor positioning systems (FIPS) [15,16], base stations are deployed as access points (APs) that transmit signals to terminal fingerprint collection devices, such as Wi-Fi [17], Bluetooth [18], and ZigBee signals [19]. The so-called fingerprint denotes arbitrary signal features that can distinguish locations, such as the received signal strength (RSS) [2,20], channel impulse response (CIR) [21], and power delay Doppler profile (PDDP). Based on the fingerprints collected by the terminal devices, two positioning phases were constructed. One is the offline fingerprint database construction phase, where the fingerprints gained by terminal devices are stored in a fingerprint database as training samples. The other one is the online positioning phase, in which one compares the fingerprint from an unknown user's position with the fingerprint database and develops a classifier to predict the user's position. Conventional FIPS consist of two phases. Nevertheless, the complex indoor environment inevitably involves unexpected dynamics, multipath effects, and other interferences, resulting in distribution differences between the source and target domains, where conventional FIPS methods can no longer guarantee positioning accuracy. Meanwhile, fingerprint collection devices may differ in offline and online phases, and even a tiny variation in the holding posture of devices tends to lead to distinct features. Additionally, a long interval between offline and online phases may cause version or quantity changes in APs, further leading to unexpected cross-domain heterogeneities of the feature spaces. Therefore, knowing how to mitigate cross-domain feature space divergence and heterogeneity is critical for accurate and robust FIPS.

Transfer learning [22], for the purpose of improving learning results through transferring knowledge from one domain to a different related domain is a promising machine learning algorithm used to mitigate distribution divergence problems for FIPS [23]. In practical FIPS scenarios, both the feature space distribution and dimension heterogeneity exist, inspiring researchers to pay more attention to heterogeneously transferred learning methods.

In the community of heterogeneous transfer learning, in order to establish the connection between source and target domains with different feature dimensions, a small number of accurately labeled samples in the target domain is normally required, which contributes to semi-supervised learning [24,25]. In the field of FIPS—especially in the online positioning phase, and different from image and text classifications—obtaining accurately labeled samples is often labor-consuming, rendering existing semi-supervised heterogeneous transfer algorithms that are not practical for FIPS in the presence of the environment or collection device changes. Furthermore, regarding transfer learning, there is a long-lasting problem with how to determine if given source samples are valuable to the target domain classification task. In some cases where the source and target domains are not strongly related, blindly using domain adaptation methods may degenerate learning properties instead, which is usually called the "negative transfer" issue. To avoid such an issue, some coefficients evaluating the correlation between two samples were developed, including the Pearson correlation coefficient (PCC) [26], Gaussian kernel function [27], and bipartite graphs [28]. Unfortunately, most of them can only evaluate correlations between homogeneous domains, motivating the extension to more general heterogeneous scenarios.

Co-occurrence data are known for their ability to capture higher-order relationships between classes, concepts, and labels [29], which are usually employed to joint data from heterogeneous domains. Conventionally, the definition of co-occurrence data involves a moment when features of both source and target samples can be captured. For instance, a picture and its corresponding text description can be treated as an image–text co-occurrence sample, which is easy to capture in the computer vision community. In FIPS, both time and feature consistencies are rather difficult to maintain, which motivates us to extend such a definition to a relatively general-sense version.

Oriented by the aforementioned problems, a heterogeneous transfer learning framework based on co-occurrence data (HTL-CD) is proposed. First, the source domain data were mapped into the target feature space, enabling us to fulfill the positioning task in the same feature space. Within our framework, the marginal and conditional probability distributions between the source and target domains were narrowed down through a maximum mean discrepancy (MMD) restriction, ensuring a higher positioning precision. Meanwhile, co-occurrence data acted as the bridge linking the source and target domains, based on which the cross-domain correlation could be calculated even with different feature spaces. Subsequently, a topology correlation restriction was imposed on each sample, such that the mapped distance between two highly related samples was particularly small, in which case, more knowledge from the source domain was transferred to the target domain. Contrarily, little knowledge transfer took place between rarely related samples, avoiding the so-called "negative transfer" problem.

The main contributions of our proposed algorithm can be summarized as

- Within the HTL-CD framework, homogeneous knowledge transfer and knowledge transfer between heterogeneous domains can be attained. To cope with the heterogeneous problem caused by the collection device heterogeneity or real-time environmental changes, a projection from the source feature space to the target feature space was considered. In addition, by imposing a MMD restriction on the mapping matrix, the cross-domain marginal and conditional distributions were both aligned, giving rise to higher positioning accuracy and robustness in uncertain environment dynamics.
- The cross-domain correlation was evaluated with a correlation coefficient, which could
  precisely reflect whether and how the given source domain is effective in knowledge
  transferring to the target domain in both homogeneous and heterogeneous cases.
  Hence, it is possible to take full advantage of a fingerprint database and avoid the
  "negative transfer" issue.
- Through extending the current definition of co-occurrence data to a general-sense version, it was dramatically easier to collect co-occurrence fingerprint samples in FIPS. With the help of general-sense co-occurrence fingerprint data, joint sources, and target domains, the correlation between two arbitrary samples can be fairly evaluated, such that even in the existence of significant cross-domain divergence, an accurate and robust positioning performance can be achieved without resorting to common features or accurately labeled samples in the target domain.

The remainder of the paper is structured as follows. Related works are reviewed in Section 2 and a problem formulation is presented in Section 3. Then, a detailed deduction of the proposed HTL-CD is presented in Section 4. The experiment environment construction and results are presented in Section 5. Finally, the conclusions of our algorithm and potential future works are discussed in Section 6. Moreover, a list of abbreviations frequently used in this article is presented in Table 1.

Notation	Meaning		
$\mathbf{D}_s, \mathbf{D}_t, \mathbf{D}_c$	source/target/co-occurrence domain		
$\mathbf{X}_{s}, \mathbf{X}_{t}, \mathbf{X}_{c}$	source/target/co-occurrence feature matrix		
$\mathbf{y}_s$ , $\mathbf{y}_t$	source/target labels		
$n_s, n_t, n_c$	number of source/target/co-occurrence samples		
$m_s, m_t, m_c$	dimension of source/target/co-occurrence samples		
$\mathbf{A}_{r}\mathbf{\hat{X}}_{s}$	mapping matrix, mapped source feature matrix		
Ν	correlation matrix		
Μ	MMD matrix		
Κ	kernelized matrix		
$1_s, 1_t$	column vectors with $m_s$ and $m_t$ elements of value 1		
α, β	topology and regularization parameters		

 Table 1. Abbreviation list.

Notation	Meaning
T	iteration times
$\ \star\ _F$	Frobenius norm of *
$(\star)^T$	transposition matrix of $\star$
$\Gamma(\star)$	fingerprint classifier

# 2. Related Work

In this section, some related works are briefly reviewed, mainly concerning classical indoor positioning algorithms and transfer learning in FIPS.

# 2.1. Classical Indoor Positioning Algorithms

#### 2.1.1. Geometric-Based Methods

Geometric-based indoor positioning typically utilizes geometric principles (such as triangulation and trilateration to calculate positions via the measured range or direction information from reference points to a target. The time-of-arrival (TOA) in [30] collects the absolute travel time of the signal from a reference node (base station, beacon, access point, and so on) to the user's equipment and computes the user's localization by trilateration. The time-difference-of-arrival (TDOA) [31] adopts the travel time difference instead, avoiding the complex synchronization task. Notice that the measurement of the range as well as the angle heavily relies on the line-of-sight (LOS) conditions, which are hardly satisfied in practical indoor applications.

#### 2.1.2. Vision-Based Methods

Due to the rapid development of computational vision techniques [32], vision-based feature extraction techniques have exhibited remarkable improvements, enabling the acquisition of target positions through several images captured by static cameras located at known positions. In [33], the authors made a comprehensive review of existing computer vision-based indoor localization methods. In [34], the authors proposed a localization framework, where a convolutional neural network (CNN) was employed to observe the type and bounding box of the target; another neural network was employed for orientation and distance calculations. Moreover, simultaneous localization and mapping (SLAM) [35] has received more attention recently; it is capable of constructing maps, positioning, and detecting indoor static objects in real time. Unfortunately, these methods tend to be environmentally vulnerable and LOS conditions are normally required, which may hinder their applications in complex dynamic indoor scenarios. Moreover, the feature extraction accuracy is positively proportional to the computational burdens of neural networks, bringing about unexpected time delays in highly accurate positioning cases and, thus, real-time positioning tasks cannot be fulfilled.

#### 2.1.3. PDR

PDR [14] recursively updates the positioning results from the last iteration based on direction and distance data given by inertial sensors, which are completely self-dependent and can operate in both LOS and non-LOS environments. However, the accumulated errors resulting from inertial sensor accuracy constraints is a fatal drawback [36], and the initial heading direction and position are required every time, which heavily limits its real-world application.

## 2.1.4. FIPS

FIPS are defined as signal features that are valuable for position classifications, such as fingerprints, RSS [2,37], and channel state information (CSI) [38]; as an example, conventional FIPS contain two phases, i.e., the offline fingerprint database construction phase and online positioning phase. In [39], by employing the fingerprint database as the training

set, the maximum likelihood position of the target could be deduced through statistical inference. Moreover, refs. [40,41] computed the similarity between the fingerprint database and testing samples via Euclidean distance, cosine similarity, or other similarity matrices, and localized targets at the fingerprint location with the highest similarity. Additionally, ref. [42] incorporated knowledge distillation into CNN-based indoor positioning systems to distill knowledge from large deep CNNs into small CNNs; lower average positioning errors could be obtained via a simplified model. The teacher-assistant framework was deployed in [43] to allow for a simple CNN indoor positioning system, where knowledge from a large pre-trained network was transferred to a small network. Furthermore, due to the pervasive wireless infrastructures empowered by 802.11 Wi-Fi protocols, Wi-Fi-based FIPS [44,45] are easy-implemented in most indoor environments without any prior location information of Wi-Fi base stations. Wi-Fi base stations are mainly designed for information exchanges instead of localization, limiting localization accuracy. Moreover, in the offline and online phases of FIPS, different terminal devices, such as IOS and Android smartphones, as well as different holding postures, result in appreciable feature fluctuations and environmental changes because of pedestrian movements, furniture changes, and other indoor dynamics, further rendering the source and target domains with distribution and dimension heterogeneities, which are the major concerns of this paper.

#### 2.2. Transfer Learning in FIPS

Transfer learning, as a different kind of machine learning method, concentrates on storing knowledge gained while solving one problem and applying it to a different (but related) problem [46]. Specifically, domain adaptation, as a classical method of transfer learning, exhibits remarkable properties in addressing cross-domain distribution differences, which is exactly one of the challenges we are aiming to conquer in FIPS. Typically, transfer component analysis (TCA) [47] and joint distribution adaptation (JDA) [48,49] are effective at narrowing down the distribution discrepancies between fingerprint databases and target samples, facilitating the subsequent high-accuracy positioning. In [50], a transfer learning-based approach was employed to learn common patterns in fingerprint data from different Wi-Fi RSS indoor positioning datasets, and smaller positioning errors were achieved. To take full advantage of finite training data, [51] used parameters of a pretrained neural network trained with the data obtained from finite difference time domain simulations. Moreover, in most practical scenarios, real-time environmental changes and collection device variations tend to cause feature space diversity, inspiring researchers to develop heterogeneous transfer learning frameworks for FIPS. In [52], a common feature space was constructed by resorting to a cross-domain mapping, allowing to adopt domain adaptation methods for further knowledge transfer. Furthermore, [53] proposed a new feature extraction scheme by retaining only the most significant predictors; they selected the most efficient feature dimensions by utilizing a hybrid-based approach to reduce the training calibration efforts. However, the aforementioned works neglected a long-lasting problem on how to determine whether (and how) the source domain knowledge is valuable for the target positioning task. In other words, the correlation between source and target domains is hardly considered in FIPS; ideally, we want more knowledge transferred between more related domains, such that "negative transfer" can be eliminated. Some correlation metrics, such as PCC [26], Gaussian kernel function [27], and bipartite graphs [28] are meant to promote effective homogeneous knowledge transfer, while the correlation calculation between heterogeneous domains is still open for FIPS.

Co-occurrence data are known for their extraordinary abilities in jointing source and target samples even from heterogeneous domains, which provide potential solutions for us to solve feature space heterogeneity. In [54], by resorting to labeled source data and unlabeled auxiliary co-occurrence data, a hedge ensemble scheme was proposed to solve the online heterogeneous transfer issue. Moreover, Xu et al. [55] developed multiple spatial pairwise local co-occurrence descriptors to improve the resolutions of fingerprint images. Note that image-based co-occurrence data are easy to obtain, while in FIPS, the rigorous

time and feature consistencies hinder the derivations of co-occurrence fingerprint samples, inspiring us to extend such definitions to general versions.

Oriented by the aforementioned problems, we propose the HTL-CD framework for FIPS, capable of narrowing down cross-domain distribution differences, capturing the degree of correlation in both homogeneous and heterogeneous cases, and facilitating better learning performances. Therefore, higher positioning precision and robustness against uncertain dynamics can be obtained for FIPS.

## 3. Problem Formulation

In the first place, we define the samples from the fingerprint database as the source domain data  $\mathcal{D}_s = \{\mathbf{X}_s, \mathbf{y}_s\} = \{\mathbf{x}_s^i, y_s^i\}_{i=1}^{n_s}$ , where  $\mathbf{x}_s^i = \begin{bmatrix} x_s^{i,1}, x_s^{i,2}, \dots, x_s^{i,m_s} \end{bmatrix}^T$  is the *i*th fingerprint sample in the database, representing the feature of the *i*th position.  $n_s$  and  $m_s$ , respectively, denotes the number of fingerprint samples and dimensions of each fingerprint. Moreover,  $y_s^i$  is the actual label of *i*th fingerprint. Similarly,  $\mathcal{D}_t = \{\mathbf{X}_t\} = \{\mathbf{x}_t^i\}_{i=1}^{n_t}$  are defined as the unlabeled target domain data to be positioned, in which  $\mathbf{x}_t^i = \begin{bmatrix} x_t^{i,1}, x_t^{i,2}, \dots, x_t^{i,m_t} \end{bmatrix}^T$  is the fingerprint of the *i*th test sample.

Traditionally, the key principle of FIPS is training a classifier  $\Gamma(\cdot)$  based on the offline source data  $\mathcal{D}_s$ , and utilizing  $\Gamma(\mathcal{D}_t)$  to achieve positions of user targets. However, this is under the assumption of the same or similar marginal probability distributions between the source and target domains, i.e.,  $P(\mathbf{X}_s) = P(\mathbf{X}_t)$ , which is unrealistic owing to real-time environmental changes. Additionally, because of the different types and holding postures of terminal devices, some sensors may not be detected in the online positioning phase, which means the source and target domains suffer from different feature dimensions and  $m_s \neq m_t$ . In such feature-missing cases, directly using traditional FIPS contrarily lead to remarkable positioning errors.

In this article, our objective was to design accurate and robust positioning FIPS using  $D_s$  and  $D_t$ , in the existence of unexpected AP variations and environmental uncertainties.

#### 4. Htl-Cd Framework

In this section, the HTL-CD will be elaborated with the overall architecture illustrated in Figure 1. Evidently, it mainly consists of a MMD restriction, co-occurrence data construction, a correlation restriction, and optimization of the objective function.



Figure 1. Overall architecture of the proposed HTL-CD.

## 4.1. MMD Restriction

Assume that there exists a mapping matrix  $\mathbf{A} \in R^{m_s \times m_t}$ , such that the source data can be mapped into the target feature space through

$$\hat{\mathbf{X}}_{s} = \mathbf{A}^{\mathrm{T}} \mathbf{X}_{s} \tag{1}$$

The subsequent positioning is completely based on the mapped source data  $\hat{\mathbf{X}}_s = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_{n_s}]$ , such that the selection of **A** is a crucial factor in the target positioning performance. Hence, the following MMD restriction is considered, with the purpose of adapting the mapped source data better to the target domain. Specifically, we hope that the marginal distribution of the mapped source data, as well as the conditional distribution, are aligned with those of the target data, such that the cross-domain distribution discrepancy can be narrowed down. Inspired by [48], the cross-domain marginal and conditional distribution dist

$$MMD(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) = \left\| \frac{1}{n_{s}} \sum_{i=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} - \frac{1}{n_{t}} \sum_{i=1}^{n_{t}} \mathbf{x}_{t,i} \right\|_{F}^{2}$$

$$MMD(\mathbf{y}_{s}|\hat{\mathbf{X}}_{s}, \mathbf{y}_{t}|\mathbf{X}_{t}) = \sum_{c=1}^{C} \left\| \frac{1}{n_{s}^{(c)}} \sum_{i=1}^{n_{s}^{(c)}} \hat{\mathbf{x}}_{s,i}^{(c)} - \frac{1}{n_{t}^{(c)}} \sum_{i=1}^{n_{t}^{(c)}} \mathbf{x}_{t,i}^{(c)} \right\|_{F}^{2}$$
(2)

with  $n_s^{(c)}$  and  $n_t^{(c)}$  notating the number of samples labeled with *c* in the source and target domain. *C* is the total number of different labels. Respectively,  $\hat{\mathbf{x}}_{s,i}^{(c)}$  and  $\mathbf{x}_{t,i}^{(c)}$  represent the *i*th sample with label *c* in the source and target domain.

For description conciseness, we denote  $\mathbf{X} = [\hat{\mathbf{X}}_s, \mathbf{X}_t]$ . Notice that  $\mathbf{y}_t$  in (2) cannot be obtained until the second iteration, inspiring us to consider the distribution in the following two cases:

*Case 1*: In the first iteration,  $\mathbf{y}_t$  is unknown and the conditional distribution cannot be derived, such that only the marginal distribution is aligned and based on Equation (2), the distribution restriction is described as follows:

$$\min_{A} \text{MMD}(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) = \min_{A} \text{Tr} \left( \mathbf{X} \begin{bmatrix} \frac{1}{n_{s}^{2}} \mathbf{1}_{s} \mathbf{1}_{s}^{\mathrm{T}} & -\frac{1}{n_{s}n_{t}} \mathbf{1}_{s} \mathbf{1}_{t}^{\mathrm{T}} \\ -\frac{1}{n_{t}n_{s}} \mathbf{1}_{t} \mathbf{1}_{s}^{\mathrm{T}} & \frac{1}{n_{t}^{2}} \mathbf{1}_{t} \mathbf{1}_{t}^{\mathrm{T}} \end{bmatrix} \mathbf{X}^{\mathrm{T}} \right)$$
(3)

*Case 2*: During the subsequent iterations, pseudo-labels for target samples were achieved via the last iteration, enabling us to minimize both marginal and conditional distribution differences, with the objective function expressed in the form of

$$\min_{A} \left[ \text{MMD}(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) + \text{MMD}(\mathbf{y}_{s} | \hat{\mathbf{X}}_{s}, \mathbf{y}_{t} | \mathbf{X}_{t}) \right] = \min_{A} \text{Tr}(\mathbf{X}\mathbf{M}\mathbf{X}^{T}) = \min_{A} \text{Tr}\left(\mathbf{X}\begin{bmatrix}\mathbf{M}_{s} & \mathbf{M}_{st}\\\mathbf{M}_{ts} & \mathbf{M}_{t}\end{bmatrix} \mathbf{X}^{T}\right)$$
(4)

$$\mathbf{M}_{s} = \mathbf{L}_{s} + \sum_{c=1}^{C} \mathbf{L}_{s}^{(c)}, \quad \mathbf{L}_{s} = \frac{1}{n_{s}^{2}} \mathbf{1}_{s} \mathbf{1}_{s}^{\mathrm{T}}, \quad \left(\mathbf{L}_{s}^{(c)}\right)_{ij}^{c} = \begin{cases} \frac{1}{\left(n_{s}^{(c)}\right)^{2}} & \mathbf{x}_{i}, \mathbf{x}_{j} \in \mathbf{X}_{s}^{(c)} \\ 0 & \text{others} \end{cases} \\
\mathbf{M}_{t} = \mathbf{L}_{t} + \sum_{c=1}^{C} \mathbf{L}_{t}^{(c)}, \quad \mathbf{L}_{t} = \frac{1}{n_{t}^{2}} \mathbf{1}_{t} \mathbf{1}_{t}^{\mathrm{T}}, \quad \left(\mathbf{L}_{t}^{(c)}\right)_{ij}^{c} = \begin{cases} \frac{1}{\left(n_{t}^{(c)}\right)^{2}} & \mathbf{x}_{i}, \mathbf{x}_{j} \in \mathbf{X}_{t}^{(c)} \\ 0 & \text{others} \end{cases} \\
\mathbf{M}_{st} = \mathbf{L}_{st} + \sum_{c=1}^{C} \mathbf{L}_{st}^{(c)}, \quad \mathbf{L}_{st} = -\frac{1}{n_{s}n_{t}} \mathbf{1}_{s} \mathbf{1}_{t}^{\mathrm{T}}, \quad \left(\mathbf{L}_{st}^{(c)}\right)_{ij}^{c} = \begin{cases} -\frac{1}{\left(n_{s}^{(c)}\right)^{2}} & \mathbf{x}_{i} \in \mathbf{X}_{s}^{(c)}, \mathbf{x}_{j} \in \mathbf{X}_{t}^{(c)} \\ 0 & \text{others} \end{cases} \\
\mathbf{M}_{st} = \mathbf{L}_{st} + \sum_{c=1}^{C} \mathbf{L}_{st}^{(c)}, \quad \mathbf{L}_{st} = -\frac{1}{n_{s}n_{t}} \mathbf{1}_{s} \mathbf{1}_{t}^{\mathrm{T}}, \quad \left(\mathbf{L}_{st}^{(c)}\right)_{ij}^{c} = \begin{cases} -\frac{1}{\left(n_{s}^{(c)}\right)^{2}} & \mathbf{x}_{i} \in \mathbf{X}_{s}^{(c)}, \mathbf{x}_{j} \in \mathbf{X}_{t}^{(c)} \\ 0 & \text{others} \end{cases} \\
\mathbf{M}_{ts} = \mathbf{L}_{ts} + \sum_{c=1}^{C} \mathbf{L}_{ts}^{(c)}, \quad \mathbf{L}_{ts} = -\frac{1}{n_{t}n_{s}} \mathbf{1}_{t} \mathbf{1}_{s}^{\mathrm{T}}, \quad \left(\mathbf{L}_{ts}^{(c)}\right)_{ij}^{c} = \begin{cases} -\frac{1}{\left(n_{t}^{(c)}\right)^{2}} & \mathbf{x}_{i} \in \mathbf{X}_{s}^{(c)}, \mathbf{x}_{j} \in \mathbf{X}_{s}^{(c)} \\ 0 & \text{others} \end{cases} \\
\mathbf{M}_{ts} = \mathbf{L}_{ts} + \sum_{c=1}^{C} \mathbf{L}_{ts}^{(c)}, \quad \mathbf{L}_{ts} = -\frac{1}{n_{t}n_{s}} \mathbf{1}_{t} \mathbf{1}_{s}^{\mathrm{T}}, \quad \left(\mathbf{L}_{ts}^{(c)}\right)_{ij}^{c} = \begin{cases} -\frac{1}{\left(n_{t}^{(c)}\right)^{2}} & \mathbf{x}_{i} \in \mathbf{X}_{s}^{(c)}, \mathbf{x}_{j} \in \mathbf{X}_{s}^{(c)} \\ 0 & \text{others} \end{cases} \\
\mathbf{M}_{ts} = \mathbf{L}_{ts} + \sum_{c=1}^{C} \mathbf{L}_{ts}^{(c)}, \quad \mathbf{L}_{ts}^{c} = -\frac{1}{n_{t}n_{s}} \mathbf{1}_{s}^{\mathrm{T}}, \quad \left(\mathbf{L}_{ts}^{(c)}\right)_{ij}^{c} = \begin{cases} -\frac{1}{\left(n_{t}^{(c)}\right)^{2}} & \mathbf{x}_{s}^{c} \in \mathbf{X}_{s}^{(c)}, \mathbf{x}_{s}^{c} \in \mathbf{X}_{s}^{(c)} \\ 0 & \text{others} \end{cases} \\
\mathbf{M}_{ts}^{c} = \mathbf{L}_{ts}^{c} = \mathbf{L}_{ts}^{c} \mathbf{1}_{ts}^{c}, \quad \mathbf{L}_{ts}^{c} = -\frac{1}{\left(n_{t}n_{s}^{c}\right)^{c}} & \mathbf{L}_{ts}^{c} = \frac{1}{\left(n_{t}^{c}\right)^{c}} & \mathbf{L}_{ts}^{c} = \frac{1}{\left(n_{t}^{c}\right$$

## 4.2. Co-Occurrence Data Definition

Co-occurrence data define a data set simultaneously containing source sample features as well as target sample features, which can be formulated as  $D_c = {X_c} = {[U; V]} =$ 

 $\{[\mathbf{u}_p; \mathbf{v}_p]\}_{p=1}^{n_c}$ , where **U** shares common features with **X**<sub>s</sub>, **V** shares common features with **X**<sub>t</sub>, and  $n_c$  is the quantity of co-occurrence samples.

In the field of FIPS, heterogeneity is a rather complex problem, resulting from the heterogeneity of distributions, terminal device types, detected sensor numbers, and other environmental variations, such that there is no reasonable and specific definition for the co-occurrence data in heterogeneous FIPS. Borrowing the co-occurrence data [29] from the computer vision community, notated by narrow-sense co-occurrences for distinction, is presented as follows.

# 4.2.1. Narrow-Sense Co-Occurrence Data

It is required that the collection devices, collection times, and labels of **U** and  $X_s$ , along with those of **V** and  $X_t$ , be strictly consistent.

Due to the different types and holding postures of terminal collection devices, some sensors may not be detected in the target positioning phases, hardly allowing us to seek a collection moment containing all of the involved APs. Consequently, collecting narrowsense co-occurrence data is unpractical in the presence of real-time environmental changes, and a more general definition of co-occurrence data should be adopted.

# 4.2.2. Wide-Sense Co-Occurrence Data

The rigorous device and time consistencies between **U** and **X**<sub>s</sub>, **V** and **X**<sub>t</sub> are no longer required, such that the co-occurrence fingerprint data  $\mathbf{X}_c = [\mathbf{U}; \mathbf{V}]$  can be achieved through following the same trajectories in offline and target online phases and combining the collected fingerprints in both phases. For instance, a fingerprint sample  $u_i$  is obtained at position  $\mu_i$  in the offline phase, then we collect the fingerprint  $v_i$  at the same position  $\mu_i$  in the online positioning phase and a co-occurrence fingerprint sample  $\eta_i = [u_i; v_i]$  is derived. Repeating this process generates a group of co-occurrence fingerprint samples, capable of solving feature space heterogeneity.

In the sequel, unless declared otherwise, the so-called co-occurrence data refer to widesense co-occurrence data, which are much easier to obtain than the traditional narrow-sense co-occurrence data.

## 4.3. Correlation Restriction

Firstly, the correlation definition between two samples is divided into homogeneous and heterogeneous cases:

*Case 1*: Correlation coefficient of two homogeneous samples **x**, **y** is directly evaluated via PPC, calculated as

$$\rho(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{k=1}^d (x_i^k - \overline{x}_i)(x_j^k - \overline{x}_j)}{\epsilon + \sqrt{\sum_{k=1}^d (x_i^k - \overline{x}_i)^2} \sqrt{\sum_{k=1}^d (x_j^k - \overline{x}_j)^2}}$$
(6)

where  $\mathbf{x}_i = [x_i^1, x_i^2, \dots, x_i^d] \in \mathbb{R}^d$  and  $\mathbf{x}_j = [x_j^1, x_j^2, \dots, x_j^d] \in \mathbb{R}^d$ .  $\overline{x}_i = \frac{1}{d} \sum_{k=1}^d x_i^k$  and  $\overline{x}_j = \frac{1}{d} \sum_{k=1}^d x_j^k$  respectively denotes the mean value of  $\mathbf{x}_i$  and  $\mathbf{y}_j$ , with  $\epsilon > 0$  being a small parameter.

Theoretically, the closer  $|\rho(\mathbf{x}_i, \mathbf{x}_j)|$  is to 1, the more strongly related they are.  $\rho(\mathbf{x}_i, \mathbf{x}_j) > 0$  means  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are positively correlated, and contrarily,  $\rho(\mathbf{x}_i, \mathbf{x}_j) < 0$  implies that  $\mathbf{x}_i$  and  $\mathbf{y}_j$  are negatively correlated.

*Case* 2: Correlation coefficient of two heterogeneous samples  $\mathbf{x}_i, \mathbf{x}_j$  is defined with the aid of  $\mathcal{D}_c = {\mathbf{X}_c} = {[\mathbf{U}; \mathbf{V}]} = {[\mathbf{u}_p; \mathbf{v}_p]}_{p=1}^{n_c}$  in the following form:

$$\mathcal{X}(\mathbf{x}_i, \mathbf{x}_j) = \max_{p=1}^{n_c} \rho(\mathbf{x}_i, \mathbf{u}_p) \rho(\mathbf{x}_j, \mathbf{v}_p)$$
(7)

where

$$\rho(\mathbf{x}_{i}, \mathbf{u}_{p}) = \frac{\sum_{k=1}^{d_{x}} (x_{i}^{k} - \overline{x}_{i})(u_{p}^{k} - \overline{u}_{p})}{\epsilon + \sqrt{\sum_{k=1}^{d_{x}} (x_{i}^{k} - \overline{x}_{i})^{2}} \sqrt{\sum_{k=1}^{d_{x}} (u_{p}^{k} - \overline{u}_{p})^{2}}} \\
\rho(\mathbf{x}_{j}, \mathbf{v}_{p}) = \frac{\sum_{i=1}^{d_{y}} (x_{j}^{k} - \overline{x}_{j})(v_{p}^{k} - \overline{v}_{p})}{\epsilon + \sqrt{\sum_{i=1}^{d_{y}} (x_{j}^{k} - \overline{x}_{j})^{2}} \sqrt{\sum_{i=1}^{d_{y}} (v_{p}^{k} - \overline{v}_{p})^{2}}}$$
(8)

in which  $\mathbf{x}_i = [x_i^1, x_i^2, \dots, x_i^{d_i}] \in \mathbb{R}^{d_i}$  and  $\mathbf{x}_j = [x_j^1, x_j^2, \dots, x_j^{d_j}] \in \mathbb{R}^{d_j}$ .  $\mathbf{u}_p = [u_p^1, u_p^2, \dots, u_p^{d_j}] \in \mathbb{R}^{d_i}$  and  $\mathbf{v}_p = [v_p^1, v_p^2, \dots, v_p^{d_j}] \in \mathbb{R}^{d_j}$  respectively shares the same dimension with  $\mathbf{x}_i$  and  $\mathbf{x}_j$ . In the field of transfer learning, there exists an inherent problem as to how effective

In the field of transfer learning, there exists an inherent problem as to how effective the source domain data are in transferring knowledge to the target domain. Aimed at addressing this problem, the following constraint is imposed on **A**, which can be expressed as

$$\min_{\mathbf{A}} \sum_{\mathbf{x}_i \in \mathbf{X}_s} \sum_{\mathbf{x}_j \in \mathbf{X}_s} \rho(\mathbf{x}_i, \mathbf{x}_j) \left\| \hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j \right\|_F^2 + \sum_{\mathbf{x}_i \in \mathbf{X}_s} \sum_{\mathbf{x}_j \in \mathbf{X}_t} 2\mathcal{X}(\mathbf{x}_i, \mathbf{x}_j) \left\| \hat{\mathbf{x}}_i - \mathbf{x}_j \right\|_F^2 + \sum_{\mathbf{x}_i \in \mathbf{X}_s} \sum_{\mathbf{x}_j \in \mathbf{X}_t} \rho(\mathbf{x}_i, \mathbf{x}_j) \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_F^2$$
(9)

Evidently, by resorting to the topology constrain in (9), the higher the correlations  $x_i$  and  $x_j$  are, the closer they will be in the mapped feature space, which is the so-called correlation restriction.

Furthermore, (9) can be rewritten in the following compact form:

$$\min_{\mathbf{A}} \operatorname{Tr} \left( \mathbf{X} \mathbf{N} \mathbf{X}^{\mathrm{T}} \right) \tag{10}$$

with

$$\mathbf{N} = \mathbf{D} - \mathbf{W} \tag{11}$$

in which  $\mathbf{W}(i, j) = w(\mathbf{x}_i, \mathbf{x}_j)$  is a correlation matrix.  $\mathbf{D} = \text{diag}(\sum_j \mathbf{W}(:, j))$  and  $\mathbf{W}(:, j)$  represents the sum of the *j*th column elements of **W**.

## 4.4. Objective Function and Iterative Optimization

In order to restrict the complexity of matrix **A** and eliminate the overfitting issue, a regularization term is added along with (4) and (10), which generates the following overall objective function:

$$\min_{\mathbf{A}} \operatorname{Tr} \left( \mathbf{X} (\mathbf{M} + \alpha \mathbf{N}) \mathbf{X}^{\mathrm{T}} \right) + \operatorname{Tr} \left( \beta \mathbf{A}^{\mathrm{T}} \mathbf{A} \right)$$
(12)

with  $\alpha$  and  $\beta$  being the design parameters, respectively, controlling the topology restriction and matrix complexities.

Let  $\mathbf{M} + \alpha \mathbf{N} = \mathbf{H}$  and  $\mathbf{H} = \begin{bmatrix} \mathbf{H}_s & \mathbf{H}_{st} \\ \mathbf{H}_{ts} & \mathbf{H}_t \end{bmatrix}$  and recalling that  $\mathbf{X} = [\hat{\mathbf{X}}_s, \mathbf{X}_t]$  the final objective function  $\mathcal{F}(\cdot)$  is derived as

$$\min_{\mathbf{A}} \mathcal{F} = \min_{\mathbf{A}} \operatorname{Tr} \left( \mathbf{A}^{\mathrm{T}} \mathbf{X}_{s} \mathbf{H}_{s} \mathbf{X}_{s}^{\mathrm{T}} \mathbf{A} + 2 \mathbf{A}^{\mathrm{T}} \mathbf{X}_{s} \mathbf{H}_{st} \mathbf{X}_{t}^{\mathrm{T}} + \mathbf{X}_{t} \mathbf{H}_{t} \mathbf{X}_{t}^{\mathrm{T}} + \beta \mathbf{A}^{\mathrm{T}} \mathbf{A} \right)$$
(13)

Setting  $\frac{\partial \mathcal{F}}{\partial \mathbf{A}} = 0$  results in

$$\frac{\partial \mathcal{F}}{\partial \mathbf{A}} = \mathbf{A}^{\mathrm{T}} \mathbf{X}_{s} \mathbf{H}_{s} \mathbf{X}_{s}^{\mathrm{T}} + \mathbf{X}_{t} \mathbf{H}_{ts} \mathbf{X}_{s}^{\mathrm{T}} + \beta \mathbf{A}^{\mathrm{T}} = 0$$
(14)

Subsequently, the analytical expression of **A** in an extreme point is calculated in the following form:

$$\mathbf{A}^{\mathrm{T}} = -\left(\mathbf{X}_{t}\mathbf{H}_{ts}\mathbf{X}_{s}^{\mathrm{T}}\right)\left(\mathbf{X}_{s}\mathbf{H}_{s}\mathbf{X}_{s}^{\mathrm{T}} + \beta\mathbf{I}\right)^{-1}$$
(15)

Employing  $\hat{\mathbf{X}}_s = \mathbf{A}^T \mathbf{X}_s$ , the source domain data can finally be mapped into the feature space; in this space, traditional machine learning methods can be applied to construct a

classifier for target positioning, e.g., K-nearest neighbor (KNN) [56] and support vector machine (SVM) [57]. A complete procedure of the proposed HTL-CD is presented in Algorithm 1.

Algorithm 1 HTL-CI	framework for	fingerprint	positioning
--------------------	---------------	-------------	-------------

**Require:**  $\mathbf{X}_{s}, \mathbf{y}_{s}, \mathbf{X}_{t}, \mathbf{X}_{c} = [\mathbf{U}, \mathbf{V}], \alpha, \beta, T$  **Ensure:**  $\mathbf{A}^{T}, \hat{\mathbf{y}}_{t}$ 1: Compute N using (11); 2: Compute M using (3) and  $\mathbf{H} = \mathbf{M} + \alpha \mathbf{N}$ ; 3: Compute  $\mathbf{A}^{T}$  using (15) and  $\hat{\mathbf{X}}_{s} = \mathbf{A}^{T}\mathbf{X}_{s}$ ; 4: Compute  $\Gamma(\cdot)$  using  $\{\hat{\mathbf{X}}_{s}, \mathbf{y}_{s}\}$  and  $\mathbf{y}_{t} = \Gamma(\mathbf{X}_{t})$ 5: while Iteration times  $\leq T$  do 6: Compute M using (4) and update  $\mathbf{H} = \mathbf{M} + \alpha \mathbf{N}$ ; 7: Update  $\mathbf{A}^{T}$  using (15) and  $\hat{\mathbf{X}}_{s} = \mathbf{A}^{T}\mathbf{X}_{s}$ ; 8: Update  $\Gamma(\cdot)$  using  $\{\hat{\mathbf{X}}_{s}, \mathbf{y}_{s}\}$  and  $\mathbf{y}_{t} = \Gamma(X_{t})$ 9: end while 10: return  $\mathbf{A}^{T}$  and  $\hat{\mathbf{y}}_{t}$ 

#### 5. Experiment Results

In this section, the positioning performance comparison between the proposed HTL-CD and existing FIPS are provided, along with the robustness analyses.

#### 5.1. Environment Description

In order to validate the robustness of our HTL-CD against the environmental changes, the experiment is conducted using UJI-DB [58], consisting of Wi-Fi RSS fingerprints collected for 25 consecutive months from the 3rd and 5th floors of Universitat Jaume I library. This environment is rather complex; it not only contains tables, chairs, benches, bookshelves, and other obstacles, it suffers from pedestrian movements, furniture changes, and other indoor dynamics, involving great challenges to our positioning tasks. Figure 2 presents an actual picture of the experiment environment, where 106 grids from 154.2 m<sup>2</sup> were respectively selected to collect fingerprint samples on the 3rd and 5th floors. Moreover, Figure 3 depicts the layout and device deployments of the experimental testbed, where the red asterisks represent the third floor's devices, and the blue asterisks represent the fifth floor's devices. In total, there were 620 Wi-Fi APs with unknown localizations.

We set the fingerprint samples from the 1st month as the source domain and samples from other months as the target domain. In Figure 4, the number of effective APs and variations compared with the 1st month are depicted to reflect the sensor changes in 25 months. Emerging APs of the *i*th month define the APs detected in the *i*th month (and were not detected in the 1st month); the missing APs represent those detected in the 1st month (and were not detected in the *i*th month). Effective APs tremendously changed, especially after the 11th month, which severely deteriorated the positioning accuracies because of different feature spaces. Moreover, the number of the same APs from the 1st month and the average positioning errors directly using KNN are presented in Figure 4, implying that the positioning errors drastically increased as effective APs decreased, i.e., traditional classification schemes, such as KNN and SVM, could not guarantee positioning accuracy due to heterogeneous cross-domain feature spaces.



Figure 2. Experimental environment.



Figure 3. Layout of the experimental testbed.



Figure 4. AP variations during the 25 months.

## 12 of 17

# 5.2. Positioning Results

The average localization error (ALE) was adopted to benchmark the accuracy capabilities for positioning algorithms, which can be formulated as

$$E = \frac{1}{n_t} \sum_{i=1}^{n_t} ||h(\hat{y}_t^i) - h(y_t^i)||_2$$
(16)

where  $h(\cdot)$  is a function transforming original label into 2D Cartesian coordinates.  $\hat{y}_t^i$  is the prediction value of the actual label  $y_t^i$ .

To thoroughly and fairly compare the positioning performance, the following two cases were considered to conduct positioning experiments using the proposed HTL-CD along with current state-of-the-art schemes, including JDA [48], balanced distribution adaptation (BDA) [59], heterogeneous daily living activity learning (HDLAL) [25], heterogeneous domain adaptation using manifold alignment (DAMA) [60], and KNN.

# Case 1: positioning with partly different features.

This case is designed to simulate a scenario with some sensors updated in the online positioning phase, i.e., the source and target domains share partially different features. Utilizing the samples from the 1st month as the source domain, the samples from the 12–25th months were the target domains to conduct our experiment. In addition, 100 co-occurrence samples were randomly selected by combining a source sample and a target sample with the same but unknown labels.

The ALEs of these algorithms are depicted in Figure 5; apparently, HTL-CD exhibited more accurate and robust positioning results. Heterogeneous methods (HDLAL and DAMA) performed even worse than the homogeneous JDA and BDA methods, mainly because they did not take full advantage of the common features. By contrast, although our algorithm did not rely on common features, it employed co-occurrence data to joint all the samples via a correlation restriction, promoting knowledge transfer between related samples, such that the "negative transfer" issue was avoided and higher positioning precision and robustness could be gained.



Figure 5. Comparison of the positioning performances with partially different APs.

#### Case 2: positioning with totally different features.

As a further illustration, an extreme circumstance was taken into account, where all the sensors were changed in the online positioning phase and the source and target domains owned totally different features. In accordance with case 1, with 2880 samples from the 1st month set as the source domain samples, and the samples from the 12th to 25th month

being target domain samples, positioning tasks were fulfilled with the help of 100 randomly chosen co-occurrence samples.

The ALEs of our proposed HTL-CD along with state-of-the-art methods are presented in Figure 6; HTL-CD exhibits the most accurate and stable positioning properties. The source domain samples of homogeneous transfer methods mainly consisted of cooccurrence samples, thus they could not provide acceptable positioning precision due to their insufficient quantity, which can be observed in Figure 6. For comparison, heterogeneous methods obviously outperformed homogeneous ones, especially by resorting to co-occurrence samples, which relate the source domain samples with target domain samples; the proposed HTL-CD exhibits the highest positioning accuracy with ALE uniformly lower than 2.9 m, further demonstrating the superiority of the proposed position algorithm.



Figure 6. Comparison of the positioning performances with totally different APs.

## 5.3. Robustness Analyses

Notice that the design parameters  $\alpha$  and  $\beta$  in (12) may affect the positioning results; this section is presented for the sensitivity analyses. Let  $\alpha$  range from 0 to 10 and  $\beta$  from 0 to 1; the corresponding ALEs are plotted in Figure 7. This figure implies that as long as  $\alpha$  and  $\beta$  are within a reasonable interval, i.e.,  $1 \le \alpha \le 8$  and  $0.2 \le \beta \le 1$ , the ALE will fluctuate slightly within an acceptable range. Contrarily, if  $\alpha$  or  $\beta$  is too small, the ALE will be much higher, which indicates that both correlation and complexity restrictions are effective in establishing accurate and robust FIPS. Hence, we can conclude that our proposed HTL-CD is not sensitive to parameters  $\alpha$  and  $\beta$ , and high-accuracy positioning does not rely on specified values of design parameters.



**Figure 7.** ALEs versus different  $\alpha$  and  $\beta$ .

On the other hand, to study the dependence of the proposed HTL-CD upon specific co-occurrence samples, 10 experiments were further conducted, each with 100 randomly selected co-occurrence samples. The other design parameters were set in accordance with the last section. The ALEs of these experimental results are depicted in Figure 8. All positioning results were satisfactory with ALEs uniformly smaller than 2.7 m, which validates the effectiveness of our HTL-CD with random co-occurrence samples.



Figure 8. Experiments with random co-occurrence samples.

## 6. Conclusions

In this article, an HTL-CD framework is proposed, aimed at solving the time-varying distribution difference and feature space heterogeneity for FIPS, which can realize accurate and robust positioning performances without repeatedly reconstructing fingerprint databases. To solve the feature space heterogeneity issue, a cross-domain projection was considered, allowing us to design FIPS in a homogeneous space. Then, a MMD restriction was imposed on the projection; both marginal and conditional distribution differences could be minimized. Wide-sense co-occurrence data were used for joint sources and target domains, which are much easier to capture than traditional narrow-sense co-occurrence data in FIPS. Hence, correlation coefficients between both homogeneous and heterogeneous domains can be calculated and based on the employed correlation restriction; more knowledge transfer occurs between more correlated samples, which take full advantage of source samples and avoid the "negative transfer" issue. Subsequently, the MMD and correlation restriction are simultaneously embedded in an overall objective function; by iterative optimizations, optimal projection and positioning results can be obtained. An experimental performance implies that the proposed HTL-CD can obtain at least 17.15% smaller ALEs than existing transfer learning-based positioning methods, which further validates the effectiveness and superiority of our algorithm. Simultaneously, through robustness analyses, the positioning results of the proposed HTL-CD do not rely on specific design parameters or deliberately selected samples. Conclusively, the proposed HTL-CD can act well in long-term temporal variations in FIPS, while its application in different environments is still limited, which will be considered in future explorations. Moreover, the proposed HTL-CD scheme inevitably requires adequate samples in the source domain, which will motivate us to pay more attention to small-scale transfer FIPS in our future works.

**Author Contributions:** Conceptualization, J.H.; Formal analysis, X.G.; Investigation, H.S.; Data curation, K.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).

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

# Appendix A

Detailed deductions of Equations (3) and (4) are given as follows:

$$\begin{split} \min_{A} \text{MMD}(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) \\ &= \min_{A} \left\| \frac{1}{n_{s}} \sum_{i=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} - \frac{1}{n_{t}} \sum_{i=1}^{n_{t}} \mathbf{x}_{t,i} \right\|_{F}^{2} \\ &= \min_{A} \text{Tr} \left[ \left( \frac{1}{n_{s}} \sum_{i=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} - \frac{1}{n_{t}} \sum_{i=1}^{n_{t}} \mathbf{x}_{t,i} \right) \left( \frac{1}{n_{s}} \sum_{i=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} - \frac{1}{n_{t}} \sum_{i=1}^{n_{t}} \mathbf{x}_{t,i} \right)^{T} \right] \\ &= \min_{A} \text{Tr} \left[ \left( \frac{1}{n_{s}} \sum_{i=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} - \frac{1}{n_{t}} \sum_{i=1}^{n_{t}} \mathbf{x}_{t,i} \right) \left( \frac{1}{n_{s}} \sum_{i=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} - \frac{1}{n_{t}} \sum_{i=1}^{n_{t}} \mathbf{x}_{t,i} \right)^{T} \right] \end{split} \tag{A1} \\ &= \min_{A} \text{Tr} \left( \frac{1}{n_{s}^{2}} \sum_{i=1}^{n_{s}} \sum_{j=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} \hat{\mathbf{x}}_{s,j}^{T} + \frac{1}{n_{t}^{2}} \sum_{i=1}^{n_{t}} \sum_{j=1}^{n_{t}} \mathbf{x}_{t,i} \mathbf{x}_{t,j}^{T} - \frac{1}{n_{s}n_{t}} \sum_{i=1}^{n_{s}} \sum_{j=1}^{n_{s}} \hat{\mathbf{x}}_{s,i} \hat{\mathbf{x}}_{s,j}^{T} \right) \\ &= \min_{A} \text{Tr} \left( \frac{1}{n_{s}^{2}} \mathbf{X}_{s} (\mathbf{1}_{s} \mathbf{1}_{s}^{T}) \mathbf{X}_{s}^{T} + \frac{1}{n_{t}^{2}} \mathbf{X}_{s} (\mathbf{1}_{t} \mathbf{1}_{t}^{T}) \mathbf{X}_{t}^{T} - \frac{1}{n_{s}n_{t}} \mathbf{X}_{s} (\mathbf{1}_{s} \mathbf{1}_{t}^{T}) \mathbf{X}_{t}^{T} - \frac{1}{n_{s}n_{t}} \mathbf{X}_{t} (\mathbf{1}_{t} \mathbf{1}_{s}^{T}) \mathbf{X}_{s}^{T} \right) \\ &= \min_{A} \text{Tr} \left( \mathbf{X} \left[ -\frac{\frac{1}{n_{s}^{2}} \mathbf{1}_{s} \mathbf{1}_{s}^{T} - \frac{1}{n_{s}n_{t}} \mathbf{1}_{s} \mathbf{1}_{t}^{T} \\ -\frac{1}{n_{t}n_{s}} \mathbf{1}_{t} \mathbf{1}_{s}^{T} \\ -\frac{1}{n_{t}n_{s}}} \mathbf{1}_{t} \mathbf{1}_{s}^{T} \\ -\frac{1}{n_{t}^{2}} \mathbf{1}_{t} \mathbf{1}_{t}^{T} \\ -\frac{1}{n_{t}^{2}} \mathbf{1}_{t}^{T} \\ -\frac{1}{n_{t}^{2}} \mathbf{1}_{t}^{T} \\ -\frac{1}{n_{t}^{2}} \mathbf{1}_$$

$$\begin{split} & \min_{A} \left[ \text{MMD}(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) + \text{MMD}(\mathbf{y}_{s} | \hat{\mathbf{X}}_{s}, \mathbf{y}_{t} | \mathbf{X}_{t}) \right] \\ &= \min_{A} \left[ \text{MMD}(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) + \sum_{c=1}^{C} \left( \frac{1}{n_{s}^{(c)}} \sum_{i=1}^{n_{s}^{(c)}} \hat{\mathbf{x}}_{s,i}^{(c)} - \frac{1}{n_{t}^{(c)}} \sum_{i=1}^{n_{t}^{(c)}} \mathbf{x}_{t,i}^{(c)} \right) \left( \frac{1}{n_{s}^{(c)}} \sum_{i=1}^{n_{s}^{(c)}} \hat{\mathbf{x}}_{s,i}^{(c)} - \frac{1}{n_{t}^{(c)}} \sum_{i=1}^{n_{t}^{(c)}} \hat{\mathbf{x}}_{s,i}^{(c)} \right)^{T} \right] \\ &= \min_{A} \left[ \text{MMD}(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) + \sum_{c=1}^{C} \left( \frac{1}{(n_{s}^{(c)})^{2}} \sum_{i=1}^{n_{s}^{(c)}} \sum_{j=1}^{n_{s}^{(c)}} \hat{\mathbf{x}}_{s,i}^{(c)} (\hat{\mathbf{x}}_{s,j}^{(c)})^{T} + \frac{1}{(n_{t}^{(c)})^{2}} \sum_{i=1}^{n_{t}^{(c)}} \sum_{i=1}^{n_{t}^{(c)}} \mathbf{x}_{t,i}^{(c)} (\mathbf{x}_{t,j}^{(c)})^{T} \right. \\ &\left. - \frac{1}{n_{s}^{(c)} n_{t}^{(c)}} \sum_{i=1}^{n_{s}^{(c)}} \hat{\mathbf{x}}_{s,i}^{(c)} (\mathbf{x}_{t,j}^{(c)})^{T} - \frac{1}{n_{s}^{(c)} n_{t}^{(c)}} \sum_{i=1}^{n_{t}^{(c)}} \mathbf{x}_{t,i}^{(c)} (\hat{\mathbf{x}}_{s,j}^{(c)})^{T} \right) \right] \\ &= \min_{A} \left[ \text{MMD}(\hat{\mathbf{X}}_{s}, \mathbf{X}_{t}) + \sum_{c=1}^{C} \left( \hat{\mathbf{X}}_{s} \mathbf{L}_{s}^{(c)} \hat{\mathbf{X}}_{s}^{T} + \hat{\mathbf{X}}_{s} \mathbf{L}_{st}^{(c)} \mathbf{X}_{s}^{T} + \mathbf{X}_{t} \mathbf{L}_{ts}^{(c)} \hat{\mathbf{X}}_{s}^{T} + \mathbf{X}_{t} \mathbf{L}_{t}^{(c)} \mathbf{X}_{s}^{T} + \mathbf{X}_{t} \mathbf{L}_{t}^{(c)} \mathbf{X}_{s}^{T} \right) \right] \\ &= \min_{A} \text{Tr} \left( \mathbf{X} \begin{bmatrix} \mathbf{M}_{s} & \mathbf{M}_{st} \\ \mathbf{M}_{ts} & \mathbf{M}_{t} \end{bmatrix} \mathbf{X}^{T} \right) \\ &= \min_{A} \text{Tr} (\mathbf{X} \mathbf{M} \mathbf{X}^{T}) \end{split}$$

# References

- 1. Lau, B.P.L.; Marakkalage, S.H.; Zhou, Y.; Hassan, N.U.; Yuen, C.; Zhang, M.; Tan, U.X. A survey of data fusion in smart city applications. *Inf. Fusion* **2019**, *52*, 357–374. [CrossRef]
- Ahmad, M.O.; Ahad, M.A.; Alam, M.A.; Siddiqui, F.; Casalino, G. Cyber-Physical Systems and Smart Cities in India: Opportunities, Issues, and Challenges. Sensors 2021, 21, 7714. [CrossRef]
- 3. Tian, S.; Yang, W.; Le Grange, J.M.; Wang, P.; Huang, W.; Ye, Z. Smart healthcare: Making medical care more intelligent. *Glob. Health J.* **2019**, *3*, 62–65. [CrossRef]
- Men, J.; Jiang, P.; Zheng, S.; Kong, Y.; Zhao, Y.; Sheng, G.; Su, N.; Zheng, S. A multi-objective emergency rescue facilities location model for catastrophic interlocking chemical accidents in chemical parks. *IEEE Trans. Intell. Transp. Syst.* 2019, 21, 4749–4761. [CrossRef]
- 5. Lygouras, E.; Santavas, N.; Taitzoglou, A.; Tarchanidis, K.; Mitropoulos, A.; Gasteratos, A. Unsupervised human detection with an embedded vision system on a fully autonomous UAV for search and rescue operations. *Sensors* **2019**, *19*, 3542. [CrossRef]

- Sirmacek, B.; Riveiro, M. Occupancy prediction using low-cost and low-resolution heat sensors for smart offices. *Sensors* 2020, 20, 5497. [CrossRef] [PubMed]
- Tekler, Z.D.; Chong, A. Occupancy prediction using deep learning approaches across multiple space types: A minimum sensing strategy. *Build. Environ.* 2022, 226, 109689. [CrossRef]
- Esrafilian-Najafabadi, M.; Haghighat, F. Occupancy-based HVAC control systems in buildings: A state-of-the-art review. *Build.* Environ. 2021, 197, 107810. [CrossRef]
- 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. [CrossRef]
- Zou, H.; Zhou, Y.; Jiang, H.; Chien, S.C.; Xie, L.; Spanos, C.J. WinLight: A Wi-Fi-based occupancy-driven lighting control system for smart building. *Energy Build.* 2018, 158, 924–938. [CrossRef]
- 11. 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]
- 12. Guvenc, I.; Chong, C.C. A survey on TOA based wireless localization and NLOS mitigation techniques. *IEEE Commun. Surv. Tutor.* **2009**, *11*, 107–124. [CrossRef]
- Charan, G.; Osman, T.; Hredzak, A.; Thawdar, N.; Alkhateeb, A. Vision-Position multi-modal beam prediction using real millimeter wave datasets. In Proceedings of the 2022 IEEE Wireless Communications and Networking Conference (WCNC), Austin, TX, USA, 10–13 April 2022; pp. 2727–2731.
- 14. Hou, X.; Bergmann, J. Pedestrian dead reckoning with wearable sensors: A systematic review. *IEEE Sens. J.* **2020**, *21*, 143–152. [CrossRef]
- 15. Dinh, T.M.T.; Duong, N.S.; Sandrasegaran, K. Smartphone-based indoor positioning using BLE iBeacon and reliable lightweight fingerprint map. *IEEE Sens. J.* 2020, 20, 10283–10294. [CrossRef]
- 16. Wang, B.; Gan, X.; Liu, X.; Yu, B.; Jia, R.; Huang, L.; Jia, H. A novel weighted KNN algorithm based on RSS similarity and position distance for Wi-Fi fingerprint positioning. *IEEE Access* 2020, *8*, 30591–30602. [CrossRef]
- 17. Yang, C.; Shao, H.R. Wi-Fi-based indoor positioning. IEEE Commun. Mag. 2015, 53, 150–157. [CrossRef]
- 18. 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]
- 19. Niu, J.; Wang, B.; Shu, L.; Duong, T.Q.; Chen, Y. ZIL: An energy-efficient indoor localization system using ZigBee radio to detect Wi-Fi fingerprints. *IEEE J. Sel. Areas Commun.* 2015, 33, 1431–1442. [CrossRef]
- Ma, Y.; Zhou, G.; Wang, S. Wi-Fi sensing with channel state information: A survey. ACM Comput. Surv. (CSUR) 2019, 52, 1–36. [CrossRef]
- Hsieh, C.H.; Chen, J.Y.; Nien, B.H. Deep learning-based indoor localization using received signal strength and channel state information. *IEEE Access* 2019, 7, 33256–33267. [CrossRef]
- 22. He, X.; Chen, Y.; Ghamisi, P. Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* 2019, *58*, 3246–3263. [CrossRef]
- 23. Wang, Y.; Lei, Y.; Zhang, Y.; Yao, L. A robust indoor localization method with calibration strategy based on joint distribution adaptation. *Wirel. Netw.* **2021**, *27*, 1739–1753. [CrossRef]
- Feng, S.; Li, B.; Yu, H.; Liu, Y.; Yang, Q. Semi-Supervised Federated Heterogeneous Transfer Learning. *Knowl.-Based Syst.* 2022, 252, 109384. [CrossRef]
- 25. Myagmar, B.; Li, J.; Kimura, S. Heterogeneous daily living activity learning through domain invariant feature subspace. *IEEE Trans. Big Data* **2020**, *7*, 922–929. [CrossRef]
- Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. In Noise Reduction in Speech Processing; Springer: Berlin, Germany, 2009; pp. 1–4.
- 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, Beijing, China, 12–16 August 2012; pp. 1–9.
- 28. Zhu, Y.; Chen, Y.; Lu, Z.; Pan, S.J.; Xue, G.R.; Yu, Y.; Yang, Q. Heterogeneous transfer learning for image classification. In Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 7–11 August 2011.
- 29. Mensink, T.; Gavves, E.; Snoek, C.G. Costa: Co-occurrence statistics for zero-shot classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 2441–2448.
- 30. Van Haute, T.; Verbeke, B.; De Poorter, E.; Moerman, I. Optimizing time-of-arrival localization solutions for challenging industrial environments. *IEEE Trans. Ind. Inform.* **2016**, *13*, 1430–1439. [CrossRef]
- Ye, X.; Rodríguez-Piñeiro, J.; Liu, Y.; Yin, X.; Pérez Yuste, A. A novel experiment-free site-specific TDOA localization performanceevaluation approach. Sensors 2020, 20, 1035. [CrossRef]
- Esteva, A.; Chou, K.; Yeung, S.; Naik, N.; Madani, A.; Mottaghi, A.; Liu, Y.; Topol, E.; Dean, J.; Socher, R. Deep learning-enabled medical computer vision. NPJ Digit. Med. 2021, 4, 1–9. [CrossRef]
- 33. Morar, A.; Moldoveanu, A.; Mocanu, I.; Moldoveanu, F.; Radoi, I.E.; Asavei, V.; Gradinaru, A.; Butean, A. A comprehensive survey of indoor localization methods based on computer vision. *Sensors* **2020**, *20*, 2641. [CrossRef]

- Hoyer, L.; Steup, C.; Mostaghim, S. A Robot Localization Framework Using CNNs for Object Detection and Pose Estimation. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bengaluru, India, 18–21 November 2018; pp. 1388–1395.
- 35. Trejos, K.; Rincón, L.; Bolaños, M.; Fallas, J.; Marín, L. 2D SLAM Algorithms Characterization, Calibration, and Comparison Considering Pose Error, Map Accuracy as Well as CPU and Memory Usage. *Sensors* **2022**, *22*, 6903. [CrossRef]
- 36. Seel, T.; Kok, M.; McGinnis, R.S. Inertial sensors applications and challenges in a nutshell. Sensors 2020, 20, 6221. [CrossRef]
- Ibrahim, M.; Torki, M.; ElNainay, M. CNN based indoor localization using RSS time-series. In Proceedings of the 2018 IEEE Symposium on Computers and Communications (ISCC), Natal, Brazil, 25–28 June 2018; pp. 01044–01049.
- Pecorella, T.; Fantacci, R.; Picano, B. Improving CSI Prediction Accuracy with Deep Echo State Networks in 5G Networks. Sensors 2020, 20, 6475. [CrossRef] [PubMed]
- He, S.; Chan, S.H.G.; Yu, L.; Liu, N. Maxlifd: Joint maximum likelihood localization fusing fingerprints and mutual distances. *IEEE Trans. Mob. Comput.* 2018, 18, 602–617. [CrossRef]
- Guo, X.; Zhu, S.; Li, L.; Hu, F.; Ansari, N. Accurate Wi-Fi localization by unsupervised fusion of extended candidate location set. *IEEE Internet Things J.* 2018, *6*, 2476–2485. [CrossRef]
- Guo, X.; Elikplim, N.R.; Ansari, N.; Li, L.; Wang, L. Robust Wi-Fi localization by fusing derivative fingerprints of RSS and multiple classifiers. *IEEE Trans. Ind. Inform.* 2019, 16, 3177–3186. [CrossRef]
- Mazlan, A.B.; Ng, Y.H.; Tan, C.K. A Fast Indoor Positioning Using a Knowledge-Distilled Convolutional Neural Network (KD-CNN). *IEEE Access* 2022, 10, 65326–65338. [CrossRef]
- 43. Aqilah, B.M.; Ng, Y.H.; Chee, K.T. Teacher-Assistant Knowledge Distillation Based Indoor Positioning System. *Sustainability* **2022**, 2022, 14652.
- 44. Qin, F.; Zuo, T.; Wang, X. Ccpos: Wi-Fi fingerprint indoor positioning system based on CDAE-CNN. *Sensors* 2021, 21, 1114. [CrossRef]
- Lee, S.; Kim, J.; Moon, N. Random forest and Wi-Fi fingerprint-based indoor location recognition system using smart watch. *Hum.-Centric Comput. Inf. Sci.* 2019, 9, 1–14. [CrossRef]
- 46. Bozinovski, S. Reminder of the first paper on transfer learning in neural networks, 1976. *Informatica* 2020, 44. 10.31449/inf.v44i3.2828. [CrossRef]
- 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]
- Long, M.; Wang, J.; Ding, G.; Sun, J.; Yu, P.S. Transfer feature learning with joint distribution adaptation. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 1–8 December 2013; pp. 2200–2207.
- 49. Sanodiya, R.K.; Yao, L. A subspace based transfer joint matching with Laplacian regularization for visual domain adaptation. *Sensors* **2020**, *20*, 4367. [CrossRef] [PubMed]
- 50. Klus, R.; Klus, L.; Talvitie, J.; Pihlajasalo, J.; Torres-Sospedra, J.; Valkama, M. Transfer Learning for Convolutional Indoor Positioning Systems. In Proceedings of the 2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Lloret de Mar, Spain, 29 November–2 December 2021; pp. 1–8.
- Korkmaz, O. Indoor Localization with Transfer Learning. In Proceedings of the 2022 30th Signal Processing and Communications Applications Conference (SIU), Safranbolu, Turkey, 15–18 May 2022; pp. 1–4.
- 52. Li, L.; Guo, X.; Zhao, M.; Li, H.; Ansari, N. TransLoc: A heterogeneous knowledge transfer framework for fingerprint-based indoor localization. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 3628–3642. [CrossRef]
- 53. Gidey, H.T.; Guo, X.; Li, L.; Zhang, Y. Heterogeneous Transfer Learning for Wi-Fi Indoor Positioning Based Hybrid Feature Selection. *Sensors* 2022, *22*, 5840. [CrossRef] [PubMed]
- 54. Yan, Y.; Wu, Q.; Tan, M.; Ng, M.K.; Min, H.; Tsang, I.W. Online heterogeneous transfer by hedge ensemble of offline and online decisions. *IEEE Trans. Neural Netw. Learn. Syst.* 2017, 29, 3252–3263. [CrossRef]
- 55. Xu, Y.; Lu, Y.; Lu, G.; Li, J.; Zhang, D. Fast pore comparison for high resolution fingerprint images based on multiple co-occurrence descriptors and local topology similarities. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *51*, 5721–5731. [CrossRef]
- 56. Peterson, L.E. K-nearest neighbor. Scholarpedia 2009, 4, 1883. [CrossRef]
- 57. Kofler, C.; Muhr, R.; Spöck, G. Classifying image stacks of specular silicon wafer back surface regions: Performance comparison of CNNs and SVMs. *Sensors* 2019, 19, 2056. [CrossRef]
- 58. Mendoza-Silva, G.M.; Richter, P.; Torres-Sospedra, J.; Lohan, E.S.; Huerta, J. Long-term Wi-Fi fingerprinting dataset for research on robust indoor positioning. *Data* **2018**, *3*, 3. [CrossRef]
- Wang, J.; Chen, Y.; Hao, S.; Feng, W.; Shen, Z. Balanced distribution adaptation for transfer learning. In Proceedings of the 2017 IEEE International Conference on Data Mining (ICDM), New Orleans, LA, USA, 18–21 November 2017; pp. 1129–1134.
- 60. Wang, C.; Mahadevan, S. Heterogeneous domain adaptation using manifold alignment. In Proceedings of the Twenty-second International Joint Conference on Artificial Intelligence, Barcelona, Spain, 16–22 July 2011; pp. 1541–1546.