Vol.2020-MBL-96 No.33
Vol.2020-UBI-67 No.33
Vol.2020-CDS-28 No.33
Vol.2020-ASD-19 No.33
2020/9/30

ドメイン敵対学習を用いた Wi-Fi CSIによるドアイベント認識

Kim Heng[†1]　尾原 和也[†2]　前川 卓也[†1]　原 隆浩[†1]　村上 友規[†3]　アベセカラ ヒランタ[†3]

Domain-Adversarial Training for Door Event Detection Using Wi-Fi Channel State Information

Abstract: Door event detection has been actively studied as it has many applications such as heating, ventilation, air conditioning control, home automation, and intrusion detection, etc. However, existing method on door event detection using Wi-Fi signals rely on a large amount of training data collected in a target environment. In this paper, we present a deep learning-based method for door event detection using domain-adversarial training to extract environment-independent features of door events from Wi-Fi CSI. It can recognize door events without employing labeled training data collected in a target environment. To achieve recognition across different environments, we leverage domain-independent features of door events, namely, differential and dynamic event features, which capture inherent changes in signal propagation caused by door events regardless to the environment. We evaluated the effectiveness of our proposed method through experiments in real environments. The experimental results demonstrated that the method can achieve the state-of-the-art performance without using labeled training data from a target environment.

Keywords: Door event detection, Wi-Fi sensing, Domain-adversarial Neural Network, Wi-Fi channel state information

1. Introduction

Context recognition techniques for indoor environments are crucial for developing ubiquitous computing systems such as elder care systems, home automation, and life logging. Dedicated sensor systems have been used to detect indoor events such as the use of everyday objects and phenomena such as ambient temperature and illumination. In particular, sensor systems for detecting events of indoor objects such as open/close doors and windows have been actively studied by researchers as they have many applications such as heating, ventilation, and air conditioning (HVAC) control, home automation, intrusion detection, and monitoring elder people. Studies on leveraging off-the-shelf sensory devices for indoor event detection have been conducted recently to reduce the installation cost of dedicated sensor systems.

For example, Mahler et al. [11] and Dissanayake et al. [4] employed disused smartphones to detect door open/close events. Other studies employed commodity Wi-Fi access points (APs) providing their pervasiveness in our daily lives. For example, Ohara et al. [12] used deep learning for detecting door open/close events based on Wi-Fi channel state information (CSI) describing how a signal propagates from a transmitter to a receiver in an environment. Overall, the majority of studies on door event detection employ supervised machine learning requiring labeled training data from a target environment. For example, Ohara et al. [12] used labeled training data containing hundreds of instances of open/close events for each door in the studied environment.

In contrast, this study proposes a new method for door event detection based on transfer learning that does not require labeled training data from a target environment. According to the proposed method, a door event detector (neural network) is trained on labeled training data from source environments in advance; it is then used to recognize door events in the target

[†1] 現在，大阪大学大学院情報科学研究科
[†2] 現在，NTT コミュニケーション科学基礎研究所
[†3] 現在，NTT アクセスサービスシステム研究所

Vol.2020-MBL-96 No.33
Vol.2020-UBI-67 No.33
Vol.2020-CDS-28 No.33
Vol.2020-ASD-19 No.33
2020/9/30

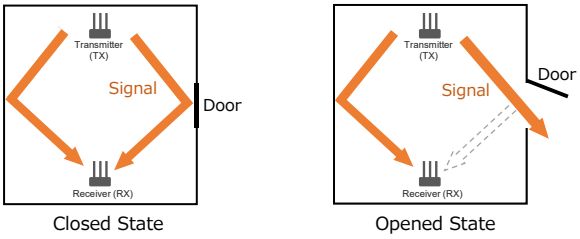


図 1: Illustration of the differences in signal propagation between door close and open states

environment. The challenge for such a transfer learning-based approach lies in the complexity of observed CSI, which contains data about various phenomena in the environment and is significantly affected by noise. To tackle this challenge, transfer learning is performed based on inherent phenomena that are commonly observed in CSI when door open/close events occur. In particular, this study extracts two types of inherent features related to door events: **differential event features** and **dynamic event features**. These two types of features are explained below in turn.

Here we consider two door states, closed and opened (Figure 1). In the closed state, signals from the transmitter are reflected by the door. In contrast, in the opened state, the signals reflected by the door in the closed state are propagated to the outside environment. Therefore, when comparing the signals received by the receiver between the closed and opened states, a significant attenuation of signals from the angle of the door can be observed in the opened state. Similarly, a gradual attenuation of signals from the angle of the door can be observed during the door opening. These differences across the door states can be treated as inherent phenomena of door events. Note that similar phenomena can be observed when other indoor events occur, e.g., movement of a person or a cabinet. In this study, a domain-adversarial neural network [5] is employed to extract inherent temporal dynamics of door events hidden in respective signal changes. As above, differential event features used in the event detector are extracted based on the difference in arriving signals between different states. In addition, the movement speed and direction of the door are computed as dynamic event features from CSI based on the Doppler shift. Although these dynamic features are recently used for recognizing indoor events such as human activities and positions [14], [15], [20], these features have not yet been used for recognizing door events. We believe that they are also useful for distinguishing between open and close events. In summary, differential and dynamic features containing inherent information of door events are extracted to accelerate transfer learning based on domain-adversarial training [5]. By combining the differential and dynamic features and domain-adversarial training, our method robustly recognizes door events even when other indoor events such as walking events occur.

The contributions of this paper are:

- We propose a new method for recognizing door events using transfer learning and Wi-Fi CSI. The proposed method does not require labeled training data from a target environment. To our knowledge, this is the first study on CSI-based door event detection that does not require labeled training data from a target environment.
- We propose to use domain-independent features of door events, namely, differential and dynamic event features, To achieve recognition across different environment.

2. Related Work

2.1 Indoor Event Detection

Small distributed sensors such as switch sensors, RFID tags, vibration sensors, and accelerometers have been used to detect events of indoor objects, including door open/close events [11]. Recently, detection systems relying on small numbers of sensors have been actively studied to reduce their installation cost. Barometric pressure sensors have been used to recognize door open/close events by detecting pressure changes caused by the events. Wu et al. [22] focused on the fact that a sharp pressure change can be observed in a building with HVAC systems when a door in the environment is opened and employed a smartphone barometer to detect the door events. Dissanayake et al. [4] employed smartphone active/passive sound sensing to detect door open/close events.

2.2 Context Recognition Based on Radio Frequencies

2.2.1 Wi-Fi Channel State Information

In a wireless network, the orthogonal frequency-division multiplexing (OFDM) systems can be modeled as

$$\boldsymbol{Y}_i = \boldsymbol{H}_i \boldsymbol{X}_i + \boldsymbol{N}_i$$

where $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ are the M_T-dimensional transmitted signal vector and M_R-dimensional received signal vector for the ith subcarrier, respectively, with M_T denoting the number of transmitting antennas and M_R denoting the number of receiving antennas in an $M_T \times M_R$ MIMO system; $\boldsymbol{N}_i$ is the M_R-dimensional noise vector, and $\boldsymbol{H}_i$ is the $M_T \times M_R$-dimensional channel matrix known as the CSI for the ith subcarrier. Let h_{mn} be element of $\boldsymbol{H}_i$. It denotes the CSI for the pair of the mth transmitting antenna and nth receiving antenna. Note that h_{mn} is a complex value represented as

$$h_{mn} = ||h_{mn}||e^{j\angle h_{mn}}$$

where $\angle h_{mn}$ denotes the phase of h_{mn}.

Vol.2020-MBL-96 No.33
Vol.2020-UBI-67 No.33
Vol.2020-CDS-28 No.33
Vol.2020-ASD-19 No.33
2020/9/30

2.2.2 Activity Recognition

Wi-Fi CSI has also been used to achieve device-free activity recognition. For example, Wang et al. [21] attempted to detect primitive actions using CSI and estimate activities comprising primitive actions using multi-class support vector machines. In [20], the authors proposed two models for quantitatively correlating CSI dynamics and human activities: a CSI-speed model that correlates CSI dynamics with the movement speed and a CSI-activity model that correlates the movement speed of different body parts with a specific activity.

Device-free fall detection for elder care support is another typical application of Wi-Fi CSI. For example, the WiFall system proposed by [6] employs the time variability and spatial diversity of CSI to detect falls in residential settings, while AntiFall [25] employs the CSI phase difference over two antennas and uses amplitude information to distinguish the fall activity from fall-like activities.

2.2.3 Door Event Detection

Ohara et al. [12] employed Wi-Fi CSI to recognize events of everyday objects, including door open/close events. A deep learning model was used to automatically extract efficient classification features. Shi et al. [18] employed FM-radio signal receivers to recognize the indoor situations "empty room," "opened door," and "walking person" based on the fact that changes in an environment impact the propagation of radio waves. Xu et al. [24] employed Wi-Fi CSI to recognize door events based on features extracted from channel frequency response (CFR) and a classifier using dynamic time warping. The authors [23] also attempted to detect door events through a wall by employing the time-reversal technique and dedicated RF devices based on CSI.

This study attempts to recognize door events *without* using labeled training data collected in a target environment.

2.2.4 Transfer Learning for CSI-based Context Recognition

To reduce the cost of collecting training data, several studies used transfer learning for CSI-based context recognition. For example, Rao et al. [16] employed transfer learning for CSI-based indoor positioning to learn feature representations as fingerprints by minimizing the distribution differences between a fingerprint database and test samples. Bu et al. [3] converted CSI data into image data and pre-trained an activity recognition model using a public image dataset for object recognition (ImageNet). Similar to our study, Jiang et al. [7] employed domain-adversarial training for activity recognition, while Wang et al. [19] employed domain-adversarial training for in-car activity recognition. In this study, inherent features of door events are extracted to accelerate domain-adversarial training, which otherwise would be difficult due to the dependence of CSI data related to door events on target environments.

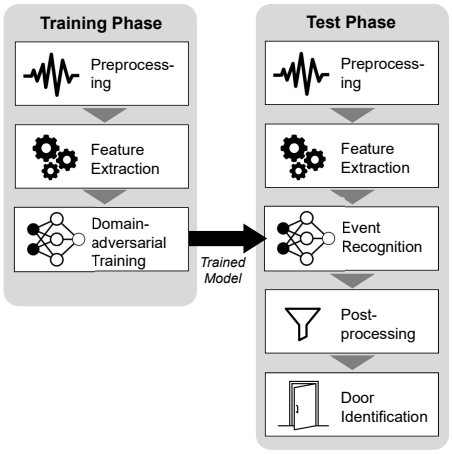


図 2: Overview of the proposed method

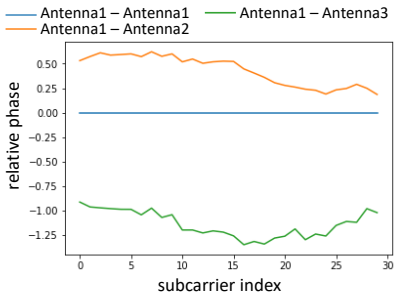


図 3: Relative phase between receiving antennas obtained using actual devices when the angle of arrival is 0 degrees.

3. Proposed method

3.1 Overview

Time-series CSI data capturing door open/close events from the receiver in the source environments were obtained and labeled in advance. Each label specified the start and end times of an event and its class (i.e., open or close class). The time-series CSI data collected in the target environment were left unlabeled.

We assume that the floor plan of each environment containing information about the positions of the receiver and doors, orientation of the receiver antenna, and type of each door (inward opening or outward opening) were available. Since door events were recognized using unlabeled training data from the target environment based on the angle of arrival (AoA) information, positional information about the receiver and doors, as well as information about the orientation of the antenna was required to distinguish between multiple doors in the environment.

Figure 2 shows an overview of the proposed method for recognizing door events in the target environment using unlabeled training data from that environment. The proposed method in-

cludes the training and test phases. In the training phase, CSI data from the source environments are preprocessed first. Dynamic and differential event features are then extracted from the preprocessed data. An event detection model (neural network) is trained on the preprocessed data with ground truth labels using domain-adversarial training. In the test phase, CSI data from the target environment are preprocessed first, and features are then extracted from the preprocessed data. The time-series of the extracted features are first recognized using the trained detection model and then processed to remove sporadic errors. Finally, each detected event is associated with a door in the target environment using AoA information.

3.2 Preprocessing

For both the training and test phases, the observed CSI data are preprocessed first since they include offsets that affect the estimation of several parameters (e.g., AoA). One of the offsets is the initial phase offset between receiving antennas. When the AoA of the Wi-Fi signal is 0 degrees, the relative phases between receiving antennas should be 0 under an ideal condition. However, the relative phases collected by actual devices significantly shift from 0 due to the initial phase offset (Figure 3).

For that, a transmitter is placed close to a receiver so that the transmitter is in front of the receive antenna array, i.e., the AoA of the direct path is 0 degrees. The initial phase offset θ_{init} can be calculated based on the collected CSI as

$$\theta_{init} = -\angle\left(\frac{1}{P}\sum_{p=1}^{P}\frac{1}{N_S}\sum_{s=1}^{N_S}\exp(j\angle(h_s(p,n,s)\overline{h}_s(p,1,s)))\right), \quad (1)$$

where $h_s(p,n,s)$ denotes the CSI collected in the static environment for the pth packet, nth receiving antenna, and sth subcarrier; P and N_S denote the number of packets and subcarriers, respectively. Then, the offset is removed as follows:

$$h_{calib}(t,n,s) = \exp(j\theta_{init})\, h_o(t,n,s), \quad (2)$$

where h_{calib} denotes the calibrated CSI and h_o denotes the CSI including the initial phase offset.

3.3 Feature Extraction

3.3.1 Dynamic Event Feature: Doppler Velocity

The Doppler velocity corresponding to the velocity of the moving object can be estimated based on phase change of CSI. Furthermore, the dynamic AoA corresponding to the direction of the moving object can be estimated using the receiving antenna array. In this section, the Doppler velocity and dynamic AoA were simultaneously estimated from preprocessed CSI data based on [15].

Let θ_i, τ_i, and v_i denote the AoA, time of flight (ToF), and Doppler velocity of the ith propagation path, respectively. Then, the CSI element of the nth receiving antenna and sth subcarrier at time t can be described as

$$h(t,n,s) = \sum_{i=1}^{P} A_i(t,n,s)\exp(-j\phi_i(t,n,s)), \quad (3)$$

where $\phi_i(t,n,s) = 2\pi\left(\phi_{i0} + (s-1)f_\delta\tau_i + \frac{f_c v_i t}{c} + \frac{f_c(n-1)d\sin\theta_i}{c}\right)$, ϕ_{i0} denotes the phase of $h(0,1,1)$ at the ith propagation path, c_f denotes the center frequency of Wi-Fi signals, and f_δ denotes the difference in frequencies between subcarriers. For simplicity, let $\boldsymbol{x} = (t,n,s)$, $\boldsymbol{\Theta} = (A_i,\tau_i,\theta_i,v_i)_{i=1}^{P}$. $\boldsymbol{\Theta}$ can be estimated using the space alternating generalized expectation maximization (SAGE) algorithm that maximizes the following log-likelihood function $\Lambda(\boldsymbol{\Theta};h)$:

$$\Lambda(\boldsymbol{\Theta};h) = -\sum_{\boldsymbol{x}}\left|h(\boldsymbol{x}) - \hat{h}(\boldsymbol{x},\boldsymbol{\Theta})\right|^2, \quad (4)$$

where $h(\boldsymbol{x})$ denotes the obtained CSI and $\hat{h}(\boldsymbol{x},\boldsymbol{\Theta})$ denotes the CSI computed from the estimated parameters $\boldsymbol{\Theta}$ using Equation 3.

The SAGE algorithm iterates over the expectation and maximization steps. Each iteration optimizes the parameter of each propagation path. The expectation step computes signal P_i for the ith propagation path as

$$\hat{P}_i(\boldsymbol{x},\hat{\boldsymbol{\Theta}}) = P_i(\boldsymbol{x},\hat{\boldsymbol{\Theta}}) + (h(\boldsymbol{x}) - \hat{h}(\boldsymbol{x},\hat{\boldsymbol{\Theta}})), \quad (5)$$

where $\hat{\boldsymbol{\Theta}}$ denotes the estimated parameters in the previous iteration. The maximization step optimizes the parameters of the ith propagation path as

$$\hat{\tau}_i' = \underset{\tau}{\operatorname{argmax}}\left\{\left|z(\tau,\hat{\theta}_i,\hat{v}_i;\hat{P}_i(\boldsymbol{x},\hat{\boldsymbol{\Theta}}))\right|\right\}, \quad (6)$$

$$\hat{\theta}_i' = \underset{\theta}{\operatorname{argmax}}\left\{\left|z(\hat{\tau}_i',\theta,\hat{v}_i;\hat{P}_i(\boldsymbol{x},\hat{\boldsymbol{\Theta}}))\right|\right\}, \quad (7)$$

$$\hat{v}_i' = \underset{v}{\operatorname{argmax}}\left\{\left|z(\hat{\tau}_i',\hat{\theta}_i',v;\hat{P}_i(\boldsymbol{x},\hat{\boldsymbol{\Theta}}))\right|\right\}, \quad (8)$$

$$\hat{A}_i' = \frac{z(\hat{\tau}_i',\hat{\theta}_i',\hat{v}_i';\hat{P}_i(\boldsymbol{x},\hat{\boldsymbol{\Theta}}))}{M_R N_S T}, \quad (9)$$

where T denotes the number of packets and

$$z(\tau,\theta,v;P_i) = \sum_{\boldsymbol{x}} P_i(\boldsymbol{x})\exp(j\phi_i(\boldsymbol{x};\tau,\theta,v)). \quad (10)$$

After convergence, v_i is estimated as the Doppler velocity, while θ_i is estimated as the dynamic AoA.

Figure 4 illustrates an example time-series of the Doppler velocity when doors are opened/closed. For example, when Door1 opens, a positive velocity can be observed. In contrast, when Door1 closes, a negative velocity can be observed.

3.3.2 Differential Event Feature: Angle of Arrival Spectrum Difference

As described in the Introduction section, the AoA of signals can change after a door event occurs. First, the AoA-ToF

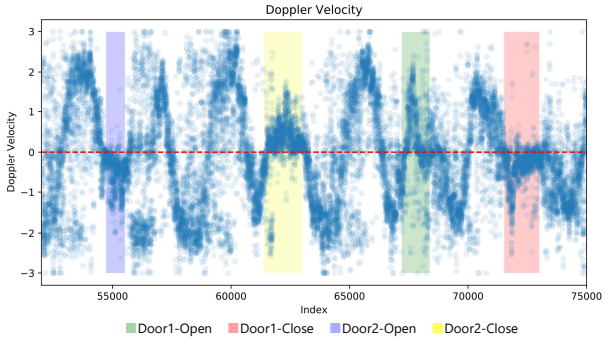


図 4: Example of the Doppler velocity

spectrum for each CSI measurement is obtained with super-resolution estimation based on [9].

The multiple signal classification (MUSIC) algorithm is employed in this study to compute the AoA-ToF spectrum from signal $\boldsymbol{X}$. First, the eigenvalues and eigenvectors of $\boldsymbol{X}\boldsymbol{X}^H$ are computed. Then, the eigenvectors whose eigenvalues are smaller than a threshold are extracted. The MUSIC algorithm computes the MUSIC spectrum $p(\theta, \tau)$ using these eigenvectors $\boldsymbol{E}_N$ as

$$p(\theta, \tau) = \frac{1}{\boldsymbol{a}^H(\theta, \tau)\boldsymbol{E}_N\boldsymbol{E}_N^H\boldsymbol{a}^H(\theta, \tau)}$$

where θ and τ denote the candidates of the AoA and ToF, respectively. When there are n' antennas and k' subcarriers, steering vector $\boldsymbol{a}(\theta, \tau)$ can be defined as

$$\boldsymbol{a}(\theta, \tau) = \left[1, \Omega, \ldots, \Omega^{(s'-1)}, \Phi, \Phi\Omega, \ldots \Phi\Omega^{(s'-1)}, \ldots, \Phi^{(n'-1)}\Omega^{(s'-1)}\right],$$

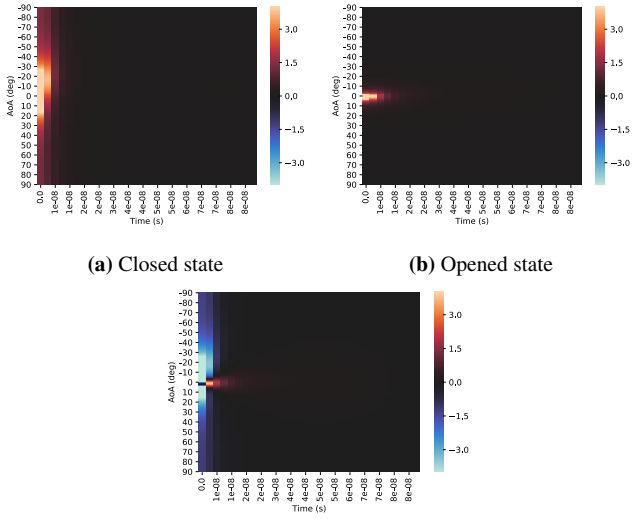
where

$$\Phi = \exp\left(-j2\pi\frac{d f_c \sin\theta}{c}\right)$$

$$\Omega = \exp\left(-j2\pi f_\delta \tau\right).$$

Here, d denotes the distance between consecutive receiving antennas, f_c denotes the center frequency of Wi-Fi signals, c denotes the speed of light, and f_δ denotes the difference in frequencies between consecutive subcarriers. By computing the value for each θ and τ, a two-dimensional AoA-ToF spectrum is obtained for each packet (Figures 5(a) and 5(b)).

The AoA spectrum differences are computed from a series of AoA-ToF spectra as follows. The reference time t_r when no door event occurs is selected as a time when the rolling standard deviation of the AoA-ToF spectra within a sliding time window is below a threshold, i.e., when the AoA-ToF spectrum is static. Next, the AoA-ToF spectrum at each time t is subtracted from that at t_r. Figures 5(a) and 5(b) illustrate examples of the AoA-ToF spectrum before and after a door open event. It can be noticed from the figures that the door event alters signal propagation. Figure 5(c) shows the difference (pixel-wise subtraction) between the spectrum illustrated



(a) Closed state **(b)** Opened state

(c) Spectrum difference

図 5: AoA-ToF spectrum and spectrum difference computed by subtracting the sepctrum of open state from that of the close state

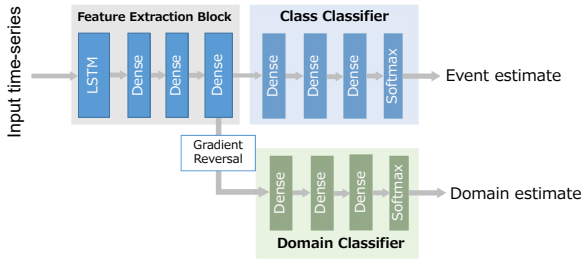


図 6: Architecture of the neural network

in Figure 5(a) and that in Figure 5(b), which is called the AoA spectrum difference in this study.

A series of AoA spectrum differences represents the dynamics of the changes in AoA from the reference time. The rolling standard deviation of the series of spectrum differences can then be computed and fed into the event detector (neural network) along with the Doppler velocity.

3.4 Domain-adversarial Training

The event detection model was trained on labeled training data from the source environments using domain-adversarial training.

The model architecture of the proposed method is illustrated in Figure 6. The input of the model is a segment of dynamic and differential event features within a sliding time window. The model has two types of outputs: domain estimate and class estimate. The domain estimate is an identifier of the estimated environment, where the input segment is observed. The class estimate is the estimated event class label of the input segment. Since the environment can have multiple doors and different types of doors (inward opening or outward opening), the input segment is first classified into none class, approaching class,

Vol.2020-MBL-96 No.33
Vol.2020-UBI-67 No.33
Vol.2020-CDS-28 No.33
Vol.2020-ASD-19 No.33
2020/9/30

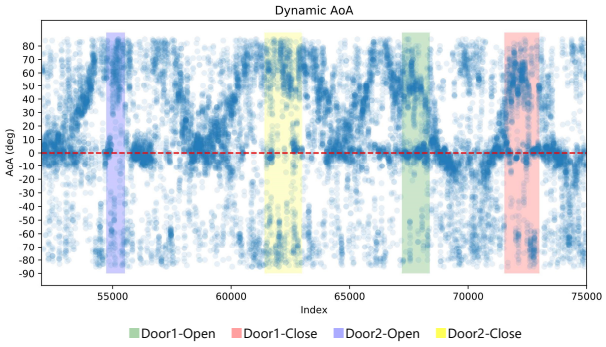


図 7: Example of dynamic AoA

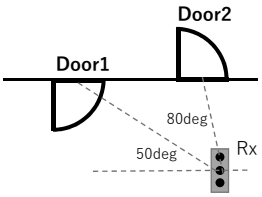


図 8: The environment where AoA data illustrated in Figure 7 were collected

or receding class to reduce the complexity of the network task. The none class means that no door events occurred. The approaching class means that an event of door movement towards the receiver occurred. The receding class means that an event of door movement away from the receiver occurred. An event is classified into approaching, when the receiver is in the room and an inward opening door opens into the room or an outward opening door closes. In contrast, an event is classified into receding, when an inward opening door in the room closes or an outward opening door opens.

The LSTM layers in the feature extraction block extract features, which are used to output the domain estimate and class estimate. The gradient reversal layer [5] is introduced to make the neural network be agnostic to domains (environments). This layer multiplies the gradient with a negative constant value during the training of the network using the back-propagation algorithm [17], enabling the LSTM layers in the feature extraction block to extract features that are indifferent to domains but sensitive to classes.

The model is trained to minimize the error of the class estimates and maximize the error of domain estimates.

3.5 Post-processing

The above procedure outputs a series of class estimates (none, approaching, or receding). To remove sporadic errors and smooth out the series of estimates, the median filter is applied to each sliding time window.

図 9: Environments of data collection

3.6 Door Identification

In this procedure, each detected door event (belonging to approaching or receding class) is associated with a door existing in the environment.

The dynamic component of the AoA (Section 3.3.1) is calculated for each event, associated with a door located in the direction of the dynamic component. Figure 7 illustrates an example of the series of dynamic AoAs computed from data collected in an environment shown in Figure 8. The figure shows the measurement of the AoA (dynamic component of AoA) for each time slice (blue dot). For the open and close events of Door2, many measurements of the AoA in the direction of around 70 degrees can be found. This means that an object movement occurred in the direction of around 70 degrees. For the open and close events of Door1, many measurements of the AoA in the direction of around 45 degrees can be found.

Let a segment between time t_n and time t_m be classified into the approaching class (or receding class). The histogram of AoA measurements within the segment is first calculated using the dynamic AoA. The AoA measurements are obtained within a segment of an event, a histogram of the AoA measurements with a bin size of five degrees is calculated, and a bin with the highest frequency is found. The angle corresponding to the found bin is an estimate of the direction of the event. Finally, the door, whose angle with respect to the receiver is the closest to the estimated direction of the event, is selected. The class label as estimated by the neural network (approaching or receding) can then be associated with the door event (open or close) by referring to the information about the door type (inward or outward).

4. Evaluation

4.1 Dataset

Sensor data were collected in seven real environments. Table 1 shows information about the environments. A Wi-Fi transmitter and a receiver with Intel 5300 NIC were installed in each environment. Figure 9 shows example frames captured by the cameras in the environments.

Throughout each session, a volunteer randomly walked in the environment to generate noises caused by a person living in the environment in our sensor data. Each session included approximately twenty door events for each door.

表 1: Experimental environments

Environment	Type	# Doors	# Sessions	Description
A	House	1 (wood)	10	Sofa is available in the environment
B	Meeting room	1 (metal+glass)	10	Desks and chairs are present
C	Meeting room	1 (metal+glass)	10	Desks and chairs are present
D	Corridor	1 (metal+glass)	8	Tables and chairs are present
E	Conference room	1 (metal)	8	Desks and chairs are present
F	Conference room	2 (metal)	10	Desks and chairs are present
G	Conference room	2 (metal)	10	Desks and chairs are present

4.2 Evaluation Methodology

The performance of the proposed method was evaluated for each environment using the leave-one-environment-out cross-validation. The following methods were employed for comparison.

- **Proposed**: The method proposed in this paper.
- **DANN-Raw**: This method performs domain-adversarial training using data from the source environments. It directly employs raw CSI data as the neural network inputs.
- **Source**: This method employs labeled data from the source environments to train an LSTM network and does not perform domain-adversarial training. It uses dynamic and differential event features as the inputs.
- **Source-Raw**: The same as **Source** except that it uses raw CSI data as the neural network inputs.
- **Target**: This method employs training data from the target environment. It uses dynamic and differential event features, and exhibits the upper bound of the event recognition accuracy.
- **Target-Raw**: The same as **Target** except that it uses raw CSI data as the neural network inputs.
- **W/o Dynamic**: This is the proposed method without considering the dynamic event features.
- **W/o Diff**: This is the proposed method without considering the differential event features.

Classification accuracy for each of the above methods was calculated based on the recognition results per window of data.

4.3 Results

4.3.1 Classification Accuracy for the Tested Methods

Figures 10 and 11 show the classification accuracy and macro-averaged F-measure for the tested methods, respectively. It can be noticed from Figures 10 and 11 that the proposed method (Proposed) achieved high performance about 86% accuracy and 83% F-measure on average despite not using any labeled data from the target environment.

According to Figure 11, the F-measure of the proposed method in Environment A is poorer than its F-measures in the other environments. This can be because Environment A is very different from the other environments (e.g., doors are made of wood).

However, introducing the differential event features improved the F-measures in Environment A as demonstrated by the results of Proposed and W/o Diff illustrated in Figure 11. The F-measure of Proposed in Environment F is also slightly poorer than its F-measures in the other environments. In contrast, the F-measure of W/o Diff in Environment F is higher than that of Proposed, indicating that the differential event features did not work well in this environment. We believe that the extracted differential event features could not capture the attenuation of signals from angles around the door because the angle between the receiver antenna and the door in Environment F was large, and it was close to the outside of the value range (sensing range) of the AoA prediction.

4.3.2 Contribution of the Proposed Event Features

It can be noticed from Figure 11 illustrating the F-measure of DANN-Raw that using raw CSI instead of the proposed features decreases the F-measure by about 40%, even when domain-adversarial training was employed. This means that it is difficult to extract environment-independent event features from raw CSI. The F-measure of Source-Raw is also very poor because features of raw CSI data significantly depend on the environment. These results indicate that the proposed features could capture environment-independent door event information.

4.3.3 Contribution of the Dynamic Event Features

It can be noticed from Figure 11 that the F-measure of W/o Dynamic is lower than that of Proposed by about 30%. This result indicates that we could successfully identify the door events from the noisy data.

4.3.4 Contribution of the Differential Event Features

It can be noticed from Figure 11 that the F-measure of W/o Diff is lower than that of Proposed by about 6%. The contribution of the differential event features is smaller than that of the dynamic event feature. However, in Environments A and B, the F-measures improved by about 10–20% by introducing the dif-

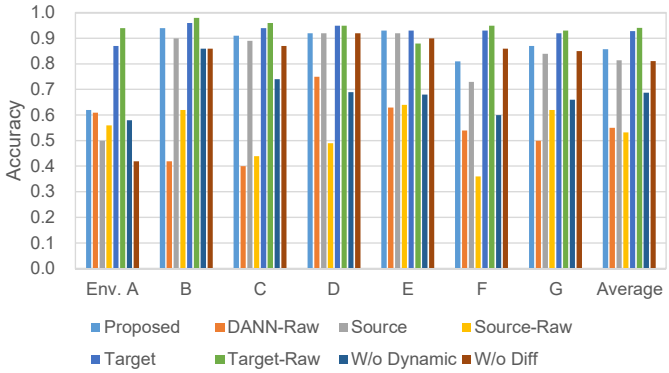


図 10: Classification accuracy scores achieved by the tested methods

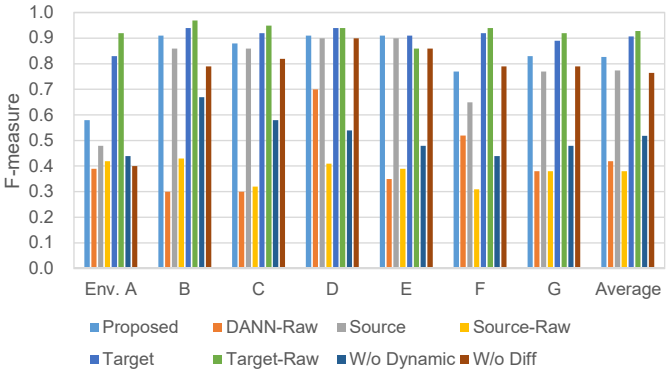


図 11: F-measure scores achieved by the tested methods

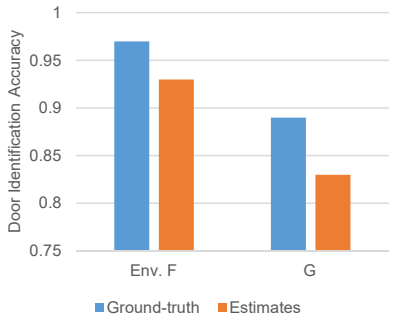


図 12: Classification accuracies for door identification when using ground-truth of door event or outputs of door event recognition (Proposed)

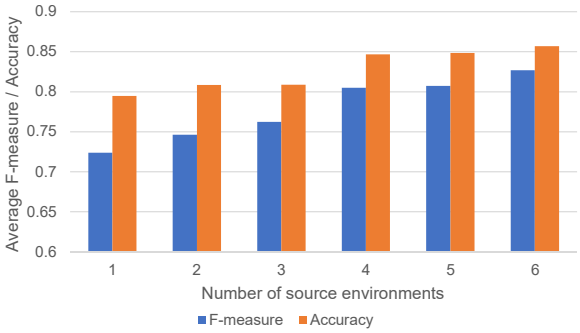


図 13: Changes in the accuracy/F-measure with the number of source environments (Proposed)

ferential event features. We consider that the differential event features are effective when source and target environments are greatly different (Environment A).

4.3.5 Contribution of Domain-adversarial Training

It can be noticed from Figure 11 that the F-measures of Source and Proposed are about 77% and 83%, respectively. Introducing the proposed domain-adversarial neural network improved the accuracy by about 6%. Significant improvements (5%–10%) can be observed in Environments A, F, and G when domain-adversarial training was employed.

4.3.6 Accuracy of Door Identification

It can be noticed from Figure 12 that the identification accuracy is very high (about 93% on average) when the ground truth data of door event recognition are employed (note that each estimate was performed for each door event). The average accuracy is about 88% when the estimates (outputs) of the proposed method are used.

4.4 Discussion on Number of Source Environments

We investigate the effect of the amount of training data on the classification accuracy. Figure 13 shows the changes in the average accuracy/F-measure with the number of source environments. (We randomly selected source environments. The computed accuracies/F-measures are the averages of three runs.) As shown in the figure, using four source environments yields the upper bound of the classification accuracy.

5. Conclusion

This study proposed a new method for detecting door events using Wi-Fi CSI data. The method does not require labeled training data from a target environment as it relies on an event detection model trained using labeled data from source environments. Since our preliminary experiments revealed that a model trained on raw CSI data does not work well in target environments, new features capturing inherent information of door events were introduced to accelerate transfer learning employed in the proposed method.

謝辞 This work is partially supported by JST CREST JPMJCR15E2, JSPS KAKENHI Grant Number JP16H06539 and JP17H04679.

参考文献

[1] Kamran Ali, Alex Xiao Liu, Wei Wang, and Muhammad Shahzad. Keystroke recognition using WiFi signals. In *the 21st Annual International Conference on Mobile Computing and Networking (MobiCom 2015)*, pages 90–102, 2015.

[2] Sheheryar Arshad, Chunhai Feng, Ruiyun Yu, and Yonghe Liu. Leveraging transfer learning in multiple human activity recognition using WiFi signal. In *2019 IEEE 20th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–10, 2019.

[3] Qirong Bu, Gang Yang, Xingxia Ming, Tuo Zhang, Jun Feng, and Jing Zhang. Deep transfer learning for gesture recognition with WiFi signals. *Personal and Ubiquitous Computing*, pages 1–12, 2020.

[4] Thilina Dissanayake, Takuya Maekawa, Daichi Amagata, and Takahiro Hara. Detecting door events using a smartphone via active sound sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(4):1–26, 2018.

[5] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, 17(1):2096–2030, 2016.

[6] Chunmei Han, Kaishun Wu, Yuxi Wang, and Lionel M Ni. WiFall: Device-free fall detection by wireless networks. In *IEEE Conference on Computer Communications (INFOCOM 2014)*, pages 271–279, 2014.

[7] Wenjun Jiang, Chenglin Miao, Fenglong Ma, Shuochao Yao, Yaqing Wang, Ye Yuan, Hongfei Xue, Chen Song, Xin Ma, Dimitrios Koutsonikolas, et al. Towards environment independent device free human activity recognition. In *the 24th Annual International Conference on Mobile Computing and Networking*, pages 289–304, 2018.

[8] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[9] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *Proceedings of the ACM Conference on Special Interest Group on Data Communication*, pages 269–282, 2015.

[10] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. Wi-Sleep: Contactless sleep monitoring via WiFi signals. In *IEEE Real-Time Systems Symposium (RTSS 2014)*, pages 346–355, 2014.

[11] Michael A Mahler, Qinghua Li, and Ang Li. SecureHouse: A home security system based on smartphone sensors. In *IEEE International Conference on Pervasive Computing and Communications (PerCom 2017)*, pages 11–20. IEEE, 2017.

[12] Kazuya Ohara, Takuya Maekawa, and Yasuyuki Matsushita. Detecting state changes of indoor everyday objects using Wi-Fi channel state information. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, 1(3):88, 2017.

[13] Shwetak N Patel, Matthew S Reynolds, and Gregory D Abowd. Detecting human movement by differential air pressure sensing in HVAC system ductwork: An exploration in infrastructure mediated sensing. In *International Conference on Pervasive Computing (Pervasive 2008)*, pages 1–18. 2008.

[14] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, Fugui He, and Tianzhang Xing. Enabling contactless detection of moving humans with dynamic speeds using CSI. *ACM Transactions on Embedded Computing Systems (TECS)*, 17(2):1–18, 2018.

[15] Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, and Yunhao Liu. Widar2.0: Passive human tracking with a single wi-fi link. In *MobiSys 2018*, pages 350–361, 2018.

[16] Xinping Rao, Zhi Li, and Yanbo Yang. Device-free passive wireless localization system with transfer deep learning method. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–17, 2020.

[17] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986.

[18] Shuyu Shi, Stephan Sigg, and Yusheng Ji. Passive detection of situations from ambient FM-radio signals. In *the 2012 ACM Conference on Ubiquitous Computing (UbiComp 2012)*, pages 1049–1053, 2012.

[19] Fangxin Wang, Jiangchuan Liu, and Wei Gong. WiCAR: WiFi-based in-car activity recognition with multi-adversarial domain adaptation. In *International Symposium on Quality of Service*, pages 1–10, 2019.

[20] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. Understanding and modeling of WiFi signal based human activity recognition. In *the 21st annual International Conference on Mobile Computing and Networking (MobiCom 2015)*, pages 65–76, 2015.

[21] Yi Wang, Xinli Jiang, Rongyu Cao, and Xiyang Wang. Robust indoor human activity recognition using wireless signals. *Sensors*, 15(7):17195–17208, 2015.

[22] Muchen Wu, Parth H Pathak, and Prasant Mohapatra. Monitoring building door events using barometer sensor in smartphones. In *the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, pages 319–323, 2015.

[23] Qinyi Xu, Yan Chen, Beibei Wang, and KJ Ray Liu. TRIEDS: Wireless events detection through the wall. *IEEE Internet of Things Journal*, 4(3):723–735, 2017.

[24] Qinyi Xu, Yi Han, Beibei Wang, Min Wu, and KJ Ray Liu. Indoor events monitoring using channel state information time series. *IEEE Internet of Things Journal*, 6(3):4977–4990, 2019.

[25] Daqing Zhang, Hao Wang, Yasha Wang, and Junyi Ma. Anti-fall: A non-intrusive and real-time fall detector leveraging CSI from commodity WiFi devices. In *International Conference on Smart homes and health Telematics (ICOST 2015)*, pages 181–193. 2015.