Indoor Events Monitoring Using Channel State Information Time Series

Qinyi Xu, *Student Member, IEEE*, Yi Han, Beibei Wang, *Senior Member, IEEE*, Min Wu, *Fellow, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—By sensing wirelessly the radio propagation environment and analyzing the channel state information (CSI), one can extend human senses beyond our traditional reach and enrich the insight into the surrounding environment and activities, with or without line-of-sight. On one hand, different indoor activities bring distinctive perturbations to wireless radio propagations. On the other hand, thanks to the nature of multipaths, indoor environmental information is contained and embedded in the wireless CSI. Since the occurrence of an indoor event lasts for a certain period of duration and repeats a similar transition pattern among different realizations, information is embedded not only in each instantaneous CSI sample, but also in how CSI changes along time, e.g., the CSI time series. Inspired by that, this paper proposes an indoor monitoring system that monitors the occurrence of different indoor events in real time with commercial WiFi devices, by exploiting the temporal information embedded in the CSI time series. Through extensive experiments, this paper studies the robustness of the proposed system to variabilities in event instances and human motion interference, and its long-term performance in a one-month test.**

***Index Terms*—Channel state information (CSI), indoor monitoring, radio analytics.**

I. Introduction

IN THE era of Internet of Things (IoT), technologies and systems have been developed to understand and decipher the surrounding environment, by answering the question of who, what, when, where, and how of everything happening. Since the past decade, billions of smart objects, also known as the "things," have been deployed around each individual. As the things form a giant network, it will be possible to comprehensively track and measure people's daily life and even monitor the entire world through IoT.

Thanks to the ubiquitous deployment of wireless radio devices and the development of emerging wireless sensing technologies, it has enabled plenty of IoT applications that utilize wireless signals, or more specifically the wireless channel state information (CSI), to perceive the information hidden in the indoor environment. *Radio analytics* has been proposed as an emerging technology that infers the propagation environment and extends the human sense over the world [1]. The feasibility of wireless passive sensing relies on multipath propagation. During the wireless transmission, wireless signals propagate through a multipath channel such that the received signal consists of copies of the transmitted signal reflected and scattered by different objects in the environment. When an object in the indoor environment moves, the resulted propagation path changes accordingly, leading to a new multipath profile.

It has gained a lot of attention for sensing with wireless signals to detect indoor events and support smart-object-based IoT applications [2], [3]. By utilizing the fact that the received radio frequency (RF) signals can be altered by the propagation environment, device-free indoor sensing systems are capable of capturing activities in the environment through the changes in the received RF signals. Existing research on wireless passive sensing can be categorized into different groups based on the features extracted from the wireless channel. To begin with, traditional wireless sensing systems mainly utilize the received signal strength (RSS) for passive monitoring applications [4]–[7] and active tracking applications [8], [9]. However, as the RSS is coarse-grained and can be easily corrupted by multipath effect, RSS-based sensing systems often require a line-of-sight (LOS) transmission, resulting in a limited accuracy in indoor activity detection.

In order to improve the accuracy and expand the application scenario of traditional wireless passive sensing, a much more informative feature, the CSI, becomes prevalent. Since the CSI typically is of high dimensions, it contains more detailed information and thus supports fine-grained classification applications. By analyzing the variations and statistics of CSI, systems have been built to detect indoor human motions [10]–[13] and small hand motions [14], [15]. Among most of those works only amplitude of the CSI was used to detect indoor activities, while the information in the phase is ignored, due to the randomness of phase distortion in the CSI. Later in [16]–[20], the first two largest eigenvalues of the CSI correlation matrix were treated as features and a support vector machine (SVM) classifier was trained to detect the presence of moving human. Although both the amplitude and phase information of the CSI was utilized in [16], it can only differentiate between the static and dynamic states in an LOS setting and the phase information was sanitized through linear fitting which has notable drawbacks. Moreover, the Doppler frequency introduced by moving objects has been extracted from CSI and one could utilize it to detect motion

Manuscript received August 27, 2018; revised December 21, 2018; accepted January 15, 2019. Date of publication January 21, 2019; date of current version June 19, 2019. *(Corresponding author: Qinyi Xu.)*

Q. Xu, B. Wang, M. Wu, and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA, and also with Origin Wireless, Inc., Greenbelt, MD 20770 USA (e-mail: qinyixu@umd.edu; bebewang@umd.edu; minwu@umd.edu; kjrliu@umd.edu).

Y. Han is with Origin Wireless, Inc., Greenbelt, MD 20770 USA (e-mail: yi.han@originwireless.net).

Digital Object Identifier 10.1109/JIOT.2019.2894332

TABLE I
COMPARISON ON RECENT WIRELESS INDOOR MONITORING SYSTEMS

Methodology	Applications	Limitations
RSSI	Object tracking [4] Activity detection [7], [24] Gesture recognition [6]	Coarse-grained feature Applicable to LOS scenario Multiple sensors deployment required
CSI	Dynamic detection [10], [16], [17], [25] Activitiy recognition [5], [5], [11], [12], [18] Intruder detection [26] Passive object tracking [27] Hand motion recognization [14], [15], [21]	Loss of CSI phase information High sounding rate ($>>$ 100 Hz) Lack of evaluation for real-time processing and detection Lack of long-term study under practical environmental changes
FMCW Radar	Human body imaging [28], [29] Passive object tracking [30]	Specially designed transmission signal High sounding rate and large antenna array Huge transmission bandwidth (over 1.7 GHz) Cannot be integrated with commercial WiFi Device calibration or specially designed hardware required

or motion directions [21]–[23]. However, in order to get an accurate Doppler shift frequency, either a high sounding rate over 1 kHz or a strict LOS transmission is required.

Another category of wireless passive sensing techniques relies on the time-of-flight (ToF) information embedded in the received signals to track changes of reflected objects for motion detection or vital sign monitoring. Due to the fact the spatial resolution of CSI is inversely proportional to the bandwidth, in order to extract the fine-grained ToF information, the ToF-based wireless sensing systems rely on either extremely large sensing bandwidths [31], [32], or specially designed frequency-modulated continuous-wave (FMCW) signals [33], [34]. Hence, those techniques cannot be implemented on off-the-shelf WiFi devices and their ability of detecting multiple indoor events has not been studied yet. In Table I, we summarize the discussion on recent wireless indoor monitoring systems, including their applications and features.

Considering the limitations of current studies we discussed above and the proliferation of demands of IoT applications in indoor monitoring, we are motivated to develop new radio analytic technique that cannot only fully utilize the information embedded in multipath channels but also support simple implementation with commercial WiFi devices. On the other hand, the occurrence of an indoor event consists of multiple states, each of which corresponds to a single multipath profile. In other words, the evolving of an indoor event is equivalent to a transition between multiple intermediate states, which uniquely determines the order and composition of a sequence of multipath profiles, as described by the CSI. Therefore, the indoor event information is embedded not only in the CSI domain but also in the temporal domain of a CSI time series.

In this paper, we propose an indoor monitoring system that monitors the occurrence of different indoor events in real time with commercial WiFi devices. Different from the existing works that only used a single multipath profile to represent an indoor event and/or state [35], [36], the system proposed in this paper exploits the temporal information embedded in the CSI time series and characterizes each indoor event as a unique dynamic transition. Because each realization of indoor event occurrences repeats a pattern of indoor state transition and lasts a period of duration, information is embedded not only in each instantaneous CSI sample but also in how CSI changes along time, e.g., the CSI time series. Instead of treating each CSI as an independent feature, the time series of CSI samples captured continuously is used for identifying and classifying different indoor events. Algorithms are designed to support real-time monitoring with high accuracy and the capability of adapting to practical environmental changes. To demonstrate the concept, we use the door opening and close in the smart home scenario as a representative set of events to study the CSI time series classification, and the technique can be generalized to other types of events. To the best of our knowledge, the proposed system is the first real-time indoor monitoring system implemented on commercial WiFi devices that has been deployed and tested in a real indoor environment for over a month without human intervention. Moreover, the proposed system does not require calibration or a high sounding rate (>100 Hz). Over the 32-day test, the proposed system successfully achieve a detection accuracy of 99.66%, which to our knowledge outperforms the existing methods.

The rest of this paper is organized as follows. System model is introduced in Section II. In Section III, we present the detailed algorithms for the proposed system, including the feature extraction algorithm, the classification algorithm for real-time monitoring, and the proposed unsupervised retraining algorithm. The performance of the proposed system is studied and evaluated through extensive experiments in Sections IV and V. We draw the conclusion in Section VI.

II. CONTINUOUS TRAJECTORY IN CSI TIME SERIES

During the wireless transmission, wireless signals propagate through a multipath channel such that the received signal consists of copies of the transmitted signal reflected and scattered by different objects in the environment. As stated in [1], multipaths can be viewed as virtual antennas, surrounding the receiver, each of which transmits an attenuated copy of the original transmit signal. A demonstration is shown in Fig. 1, where blue lines mark the reflected/scattered paths and the red line denotes the LOS path. In other words, what contains in the CSI are the characteristics of all multipaths. When an object moves inside the indoor environment, multipaths reflected and scattered by the object are changing accordingly, which can be viewed as the virtual antennas are moving. In the following, we study how the CSI, as represented by the channel frequency

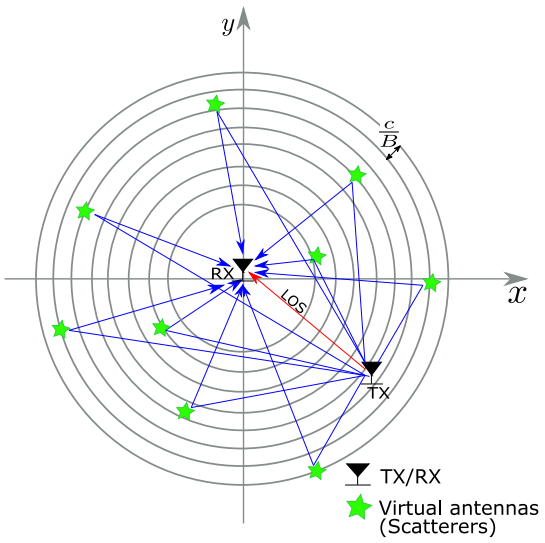


Fig. 1. Illustration of virtual antennas [37].

response (CFR), is affected by a moving object inside the propagation environment. Each object can be viewed as a collection of multiple scatterers, i.e., virtual antennas. Let us start from the case where only one virtual antenna moves while others are static. The corresponding CSI can be decomposed as

$$h_F(l,t) = \alpha_\Delta(t) e^{-j2\pi\tau_\Delta(t) f_l} + \sum_k \alpha_k e^{-j2\pi\tau_k f_l}, \quad l = 0, 1, \ldots, L-1 \tag{1}$$

where $h_F(l,t)$ is the time-varying CFR coefficient on subcarrier l with center frequency being f_l, $\alpha_\Delta(t)$ and $\tau_\Delta(t)$ denote the multipath coefficient and ToF associated to the moving antenna at current time instance t.

When the virtual antenna is moving, a sequence of $(\alpha_\Delta(t), \tau_\Delta(t))$ is uniquely associated to the moving path. $|\alpha_\Delta(t)| > 0$ and $\tau_\Delta(t) > 0$. Hence, the CFR coefficient $h_F(l,t)$ is determined by the moving path of the virtual antenna. Let $\mathbf{h}_F(t)$ denote the CFR vector as $\mathbf{h}_F(t) = [h_F(0,t), \ldots, h_F(L-1,t)]$. Because each indoor event S_i involves a set of moving virtual antennas, each S_i uniquely determines a sequence of CFR $\mathbf{h}_{F,i}(t)$'s. In other words, with the help of multipath information in $\mathbf{h}_{F,i}(t)$'s, current indoor state can be deciphered by finding out which event is happening. Ultimately, changes introduced by human activities and moving objects can be extracted from the CSI and recognized through wireless sensing.

The trajectory of a moving virtual antenna in the physical space corresponds to a continuous logical trajectory in the TR space, represented by a time sequence of CSI. Recall the CSI definition in (1) at time instance t, then we can have the CFR at time instance $t+\delta t$ ($\delta t > 0$) as

$$h_F(l,t+\delta t) = \alpha_\Delta(t+\delta t) e^{-j2\pi\tau_\Delta(t+\delta t) f_l} + \sum_k \alpha_k e^{-j2\pi\tau_k f_l}. \tag{2}$$

Assuming $\alpha_\Delta(t+\delta t) = \alpha_\Delta(t)$ during short time period, the difference between $h_F(l,t+\Delta t)$ and $h_F(l,t)$ on subcarrier l becomes

$$\begin{aligned} &|h_F(l,t+\delta t) - h_F(l,t)| \\ &\le \left|\alpha_\Delta(t)||e^{-j2\pi f_l \tau_\Delta(t)}\left(e^{-j2\pi\frac{f_l}{c}\int_t^{t+\delta t} v(u)du} - 1\right)\right| \\ &= \left|\alpha_\Delta(t)||e^{-j2\pi\frac{f_l}{c}\int_t^{t+\delta t} v(u)du} - 1\right| \end{aligned} \tag{3}$$

where $v(u)$ is the time-varying moving speed of the moving virtual antenna with respect to the TX-RX link, $\int_t^{t+\delta t} v(u)du$ represents path length change between time interval $(t, t+\delta t)$ of the propagation path associated to the virtual antenna. Furthermore, $|e^{-j2\pi(f_l/c)\int_t^{t+\delta t} v(u)du} - 1| = 0$, if and only if $(f_l/c)\int_t^{t+\delta t} v(u)du \in \mathbb{Z}$ where $\mathbb{Z}$ is the set of integer.

In other words, for any $\epsilon > 0$, we have

$$\left|e^{-j2\pi\frac{f_l}{c}\int_t^{t+\delta t} v(u)du} - 1\right| < \epsilon/|\alpha_\Delta(t)|$$
$$\forall \delta t \in \left\{\delta t \Big| \int_t^{t+\delta t} v(u)du = \frac{c}{f_l}k, k \in \mathbb{Z}\right\}. \tag{4}$$

Here, (c/f_l) is the wave length of EM waves under center frequency f_l, which is equal to 6 cm in WiFi 5G band. Hence, $\forall\epsilon > 0$, $\exists\delta > 0$ such that

$$|h_F(l,t+\delta t) - h_F(l,t)| < \epsilon \text{ and } \delta > \delta t \tag{5}$$

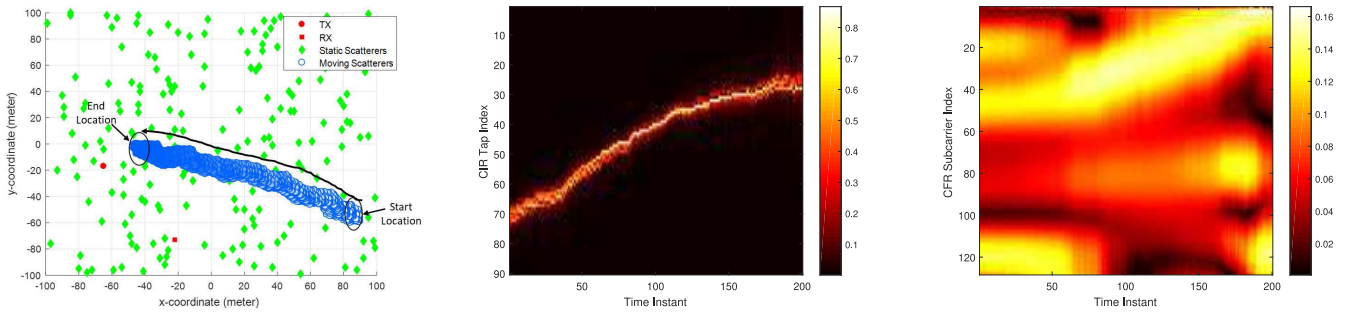
which proves the continuity of $h_F(l,t)$ on t.

Therefore, when a virtual antenna is moving, the corresponding CSI time series changes continuously. The result can be easily extended to the case of multiple moving virtual antennas, i.e., the changes introduced to the CSI by one or multiple moving objects indoors are continuous in time. As stated in [36] and [38], each indoor location or an indoor state can be viewed as a unique point in the TR space which is represented by a multipath profile. Since each position of a moving virtual antenna uniquely relates to a unique multipath profile (as described by CFR), the moving trajectory in the physical space corresponds to a unique continuous logical trajectory in the TR space.

A simulation result that studies the continuous change in the CSI time series due to moving scatterers is shown in Fig. 2. In the simulation, 200 static scatters are uniformly distributed in the area surrounding the transmitter (TX) and the receiver (RX), whose locations are marked in Fig. 2(a). The TX and the RX are transmitting under a 80-MHz bandwidth at the 5-GHz carrier frequency. A group of 50 moving scatterers located within a circle of a 5-m radius are moving under a random time-varying speed, and its final trajectory is highlighted by the color-changing blue circles in Fig. 2(a). In Fig. 2(b) and (c), the changes introduced to the CSI is plotted along time in the form of the channel impulse response (CIR) and the CFR, respectively. For both representations of the CSI, the changes due to multiple moving virtual antennas are continuous and smooth, which validates the above derivation.

A. Multiantenna Diversity

MIMO transmission introduces a large number of degrees of freedom delivered through spatial diversity for RF sensing. Suppose there is a number of $|S|$ indoor events to be monitored and let $\mathbf{h}_{F,i}^{(m,n)}[k]$ denote the kth complex-valued CSI



(a) (b) (c)

Fig. 2. Simulation of changes in the CSI time series due to moving scatterers. (a) Simulation setup: scatterers. (b) Absolute changes in the CIR time series introduced by moving scatterers. (c) Absolute changes in the CFR time series introduced by moving scatterers.

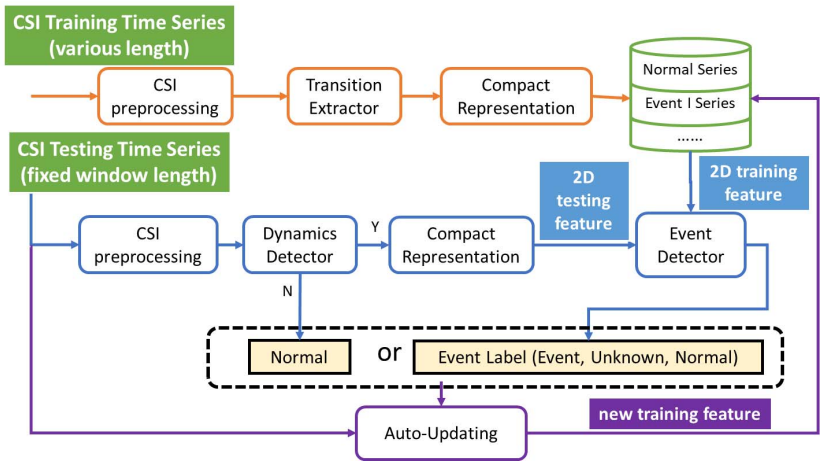


Fig. 3. System diagram.

vector, representing CFR, measured on the link between the mth TX antenna and the nth RX antenna during event $S_i \in S$. $\mathbf{h}_{F,i}^{(m,n)}[k]$ is captured at time instance kT_s, with T_s being the channel probing interval. To fully utilize the spatial diversity, we concatenate CSI vectors from different links into a single column vector as the augmented CSI by

$$\mathbf{H}_{F,i}[k] = \left[\mathbf{h}_{F,i}^{(1,1)}[k]^{\mathrm{T}}, \ldots, \mathbf{h}_{F,i}^{(N_{\mathrm{TX}},N_{\mathrm{RX}})}[k]^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{6}$$

Here, $\mathbf{H}_{F,i}[k]$ is a complex-valued column vector of length $L \times N_{\mathrm{TX}} \times N_{\mathrm{RX}}$, L denotes the number of accessible subcarriers, the superscript $\{\cdot\}^{\mathrm{T}}$ represents the transpose operator, and N_{TX} and N_{RX} denote the number of TX and RX antennas, respectively.

A real-valued waveform vector $\mathbf{G}_i[k]$ is generated by concatenating the real and imaginary part of the obtained augmented CSI $\mathbf{H}_{F,i}[k]$, i.e.,

$$\mathbf{G}_i[k] = \left[\Re\left\{\mathbf{H}_{F,i}[k]^{\mathrm{T}}\right\}, \ \Im\left\{\mathbf{H}_{F,i}[k]^{\mathrm{T}}\right\}\right]^{\mathrm{T}} \tag{7}$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are operations to take the real and imaginary part of a complex value.

Even though information on all transmission links is included in $\mathbf{G}_i[k]$'s, the dimension of feature increases dramatically and makes the classification more difficult. In this paper, we propose a feature extraction algorithm that performs refinement and dimensionality reduction on $\mathbf{G}_i[k]$'s.

III. System Design

In this section, the design of the proposed indoor monitoring system is presented. During the training phase, feature extraction algorithms are designed to refine the most distinct and representative sequence of CSI from the entire time series as the training template. We also adopt principle component analysis (PCA) to remove the correlation among different subcarriers and links, and to reduce the noise, in pursuit of a compact representation for training series. Real-time monitoring faces several practical challenges, including unknown start and end point of event occurrence, variabilities in event instances, accurate detection with low latency. To address those challenges, we propose in this paper a modified classifier based on the k-nearest-neighbor (kNN) to overcome the perturbation and divergence in the real-time measurements. The similarity between training and testing feature series can be evaluated through either Euclidean distance or dynamic time warping (DTW).

Long-term robustness is another challenge in a real-time indoor monitoring system, due to the inevitable and unpredictable changes in the environment along time. In this paper, an automated unsupervised retraining algorithm is designed for the proposed system that guarantees high accuracy against environmental changes. The system diagram is illustrated in Fig. 3 where through the proposed feature extraction algorithm, effective features will be extracted from the raw training

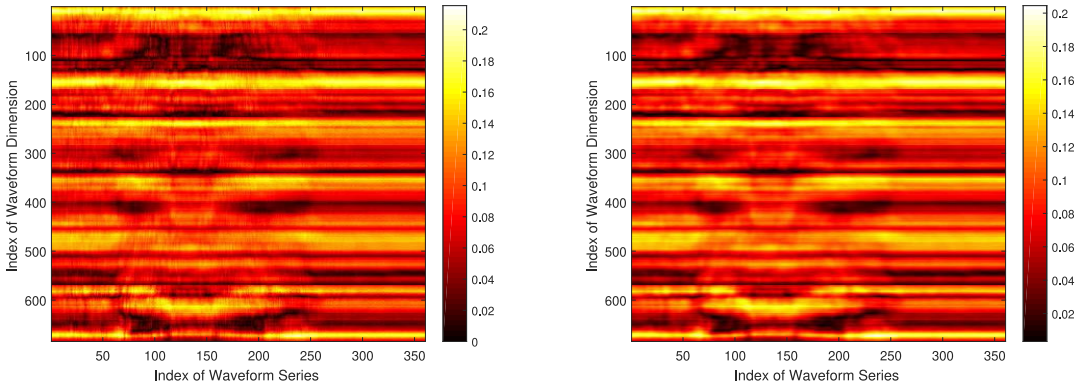


Fig. 4. Waveform spectrogram: absolute value of waveform series before and after low pass filtering. (a) Before Gaussian filtering. (b) After Gaussian filtering.

CSI time series collected during event occurrences and stored in the training database. On the other hand, in the online monitoring phase, the system first judges if the environment is dynamic based on the incoming testing CSI series after preprocessing. Then an event detector is applied to the testing feature generated from the raw CSI series to determine which event is happening, if the proposed dynamic detector detects dynamics in the environment. The proposed auto-updating algorithm works to unsupervisedly gather new candidate training sequence from the testing time series. The detailed algorithm design is discussed in the following.

A. Algorithms for Feature Extraction

In this part, we introduce the proposed algorithm that refines the measured CSI time series and extracts distinct features for all indoor events of interest during the training phase.

1) Refinement of CSI Time Series: The essential part of the proposed algorithm is to extract the most representative segment in the CSI time series captured during the occurrence of each indoor event for building a good classifier later.

Low-Pass Filtering: CSI measurements provided directly by commercial WiFi devices are often inherently noisy, due to thermal noise, noise from analog-to-digital converters, and changes in transmit power and rates. To make the measured CSI training sequences helpful and useful in representing different indoor events, the noise must be first removed from the CSI time series. In the proposed system, a Gaussian filter with length $1/3T_s$ is applied to the waveform sequence $\mathbf{G}_i[k]$'s for each event S_i, where T_s is the channel probing interval. The Gaussian filter provides a weighted averaging smoothing approach where the central data points are given with more weights and the neighbors have fewer weights. A large window length of Gaussian smoothing window results in a greater degree of filtering and a greater amount of noise reduction. However, a larger filter length also degrades the detailed information in a time series. With a length of $1/3T_s$, the moving average for the center point will only consider the data points fall into the radius of $1/6T_s$ from the center, which in our case corresponds to a duration of 0.33 s. Considering the moving object has a speed of 1 m/s, 0.33 s is a reasonable duration for the CSI perturbed by the moving object to be highly correlated and can be used to smooth out the noise. Hence, in this paper, we adopt the window length of $1/3T_s$ to filter out noise in the CSI time series without sacrificing too much detailed CSI transition information.

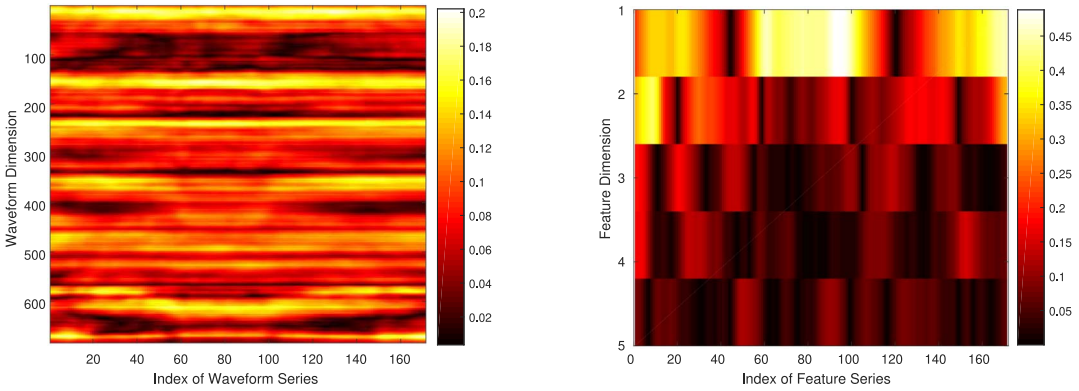
An example of waveform series before and after passing through the Gaussian filter is shown in Fig. 4, where the waveform series is measured during a door open/closed event. Compared with Fig. 4(a), the waveform series in Fig. 4(b) exhibits a much smoother transition pattern.

In the training phase, the CSI time series received at the RX may capture some indoor status similar to other indoor events at the beginning and the end part of the series. Resembling CSI subsequences, captured from different indoor events, introduce ambiguity into pattern matching and degrade the classification performance. On the other hand, from our observation over real data and the channel model in (1), information among different subcarriers and links are highly correlated. Therefore, it is necessary for applying PCA over the training waveform series to generate a compact representation, given the high dimension of data. In order to learn an efficient PCA projection matrix, it is important to keep only the dynamic and distinct transition pattern of the waveform series, and discard all the static part. As shown in Fig. 4(b), it is clear that the significant segment in $\mathbf{G}_i[k]$'s approximately starts from the 70th sample to the 240th sample, while others only contain useless static information.

To address that, a waveform extraction algorithm is proposed to track the change in waveform series and extract the most representative and dynamic segment. Taking into consideration that different links may capture different environmental information, the proposed waveform extraction evaluates link-wise dynamics. To do so, the waveform $\mathbf{G}_i[k]$ is first decomposed into $N_{\mathrm{TX}} \times N_{\mathrm{RX}}$ sequences, and can be rewritten as

$$\mathbf{G}_i[k] = \left[\mathbf{G}_i^{(1,1)}[k]^{\mathrm{T}}, \ldots, \mathbf{G}_i^{(m,n)}[k]^{\mathrm{T}}, \ldots, \mathbf{G}_i^{(N_{\mathrm{TX}},N_{\mathrm{RX}})}[k]^{\mathrm{T}}\right]^{\mathrm{T}}, \quad k = 1, 2, \ldots, M \tag{8}$$

where M is the number of samples in the time series, and each $\mathbf{G}_i^{(m,n)}[k]$ is of dimension $2L \times 1$ with L being the number of



(a) (b)

Fig. 5. Waveform spectrogram: absolute value of waveform series before and after PCA. (a) Waveform series after feature extraction. (b) Feature series after PCA.

Algorithm 1 Waveform Extraction

Input: $\mathbf{G}_i^{(m,n)}[k],\ k = 1, 2, \ldots, M, \forall(m, n)$ captured for event S_i as defined in (9) after Gaussian filtering.

Output: index $k_{s,i}$ and $k_{e,i}$ s.t. $\mathbf{G}_i[k],\ k_{s,i} \le k \le k_{e,i}$ only contains significant variations

1: Step I: Calculate $D_\delta^{(m,n)}[k] = F\left\{\left|\left(\mathbf{G}_i^{(m,n)}[k+\delta] - \mathbf{G}_i^{(m,n)}[k]\right) \cdot / \mathbf{G}_i^{(m,n)}[k]\right|\right\}, k = 1, 2, \ldots, M-\delta$, with "$\cdot/$" denoting element-wise division between vectors and $F\{\cdot\}$ being the function to take median value among all elements in a vector.

2: Step II: Obtain the best link among all $(m, n) \in \{1 \le m \le N_{TX}, 1 \le n \le N_{RX}\}$ by $(m^*, n^*) = \arg\max_{(m,n)} \sum_k^{M-\delta} D_\delta^{(m,n)}[k]$.

3: Step III: $k_{s,i} = \arg\min_k \left\{k | D_\delta^{(m^*,n^*)}[k] > \gamma\right\}$ and $k_{e,i} = \arg\max_k \left\{k | D_\delta^{(m^*,n^*)}[k] > \gamma\right\}$, with an empirical threshold γ.

accessible subcarriers on one link, given by

$$\mathbf{G}_i^{(m,n)}[k] = \left[\Re\left\{\mathbf{h}_{F,i}^{(m,n)}[k]^{\mathrm{T}}\right\},\ \Im\left\{\mathbf{h}_{F,i}^{(m,n)}[k]^{\mathrm{T}}\right\}\right]^{\mathrm{T}}. \tag{9}$$

Taking into consideration that due to spatial diversity, different links in an MIMO transmission may observe different multipath changes introduced by the same event. Hence, the proposed feature extraction algorithm evaluate transition dynamics on each link. The details are described as in Algorithm 1. An example of applying feature extraction algorithm onto the waveform series in Fig. 4 is shown in Fig. 5(a), where static parts have been discarded.

2) Denoising and Compact Representation: Unfortunately, only applying a Gaussian filter to the incoming waveform series does not yield an effective and efficient denoising outcome. Moreover, as we discussed in the previous section, the channel information on all subcarriers are highly correlated. Based on that, we proposed to apply PCA for the purpose of noise removal, de-correlation, and dimension reduction. Also, PCA is applied to waveform vectors of all indoor events after the process in Section III-A1, in order to seek an efficient feature representation that amplifies the distinction among waveforms. The details are as follows.

Let Ω_{all} denote the super waveform matrix generated by

$$\Omega_{\text{all}} = \left[\mathbf{G}_1\left[k_{s,1} : k_{e,1}\right], \ldots, \mathbf{G}_i\left[k_{s,i} : k_{e,i}\right], \ldots, \mathbf{G}_{|S|} \times \left[k_{s,|S|} : k_{e,|S|}\right]\right] \tag{10}$$

where $\mathbf{G}_i[k_{s,i} : k_{e,i}]$ denotes the waveform series after feature extraction for event S_i and Ω_{all} has a dimension of $2L \times K_{|S|}$. The mean waveform vector $\overline{\mathbf{G}}$ can be obtained by

$$\overline{\mathbf{G}} = \frac{1}{K_{|S|}} \sum_{k=1}^{K_{|S|}} \Omega_{\text{all}}[k]. \tag{11}$$

The PCA projection is learned with the correlation matrix of $(\Omega_{\text{all}}[k] - \overline{\mathbf{G}})$'s. By taking out the mean waveform $\overline{\mathbf{G}}$, we anticipate the impact of environmental background information is mitigated. Meanwhile, the projection matrix Φ of dimension $p_c \times P$ is obtained as the collection of normalized eigenvectors of the correlation matrix of $(\Omega_{\text{all}}[k] - \overline{\mathbf{G}})$'s, and p_c represents the number of principal components (PCs) to be kept. In practice, the value of p_c is selected by picking the first several largest eigenvalues that contain over 80% of the total energy among all eigenvalues. Since only the first few PCs are considered, the PCA can be computed efficiently through thin-SVD.

Then, for each event S_i, the final feature vector $\mathbf{Z}_i[k]$ can be obtained by

$$\mathbf{Z}_i[k] = \Phi \times \left(\mathbf{G}_i[k] - \overline{\mathbf{G}}\right) \quad \forall k \tag{12}$$

where the projected feature vector $\mathbf{Z}_i[k]$ is now of length p_c. An example of the comparison between $\mathbf{G}_i[k]$'s and $\mathbf{Z}_i[k]$'s is plotted in Fig. 5, where the projected feature series $\mathbf{Z}_i[k]$'s exhibits a significant variation among all PC dimensions while changes in original waveform series $\mathbf{G}_i[k]$'s is too small and too diffused to be observed.

After walking through all the preprocessing algorithms proposed in Section III-A, the final feature waveform $\mathbf{Z}_i[k]$'s will be stored in the training database for all event S_i, as the reference for the real-time monitoring.

B. Algorithms for Real-Time Monitoring

In this section, we present algorithms designed for real-time monitoring phase, which addresses variabilities in event

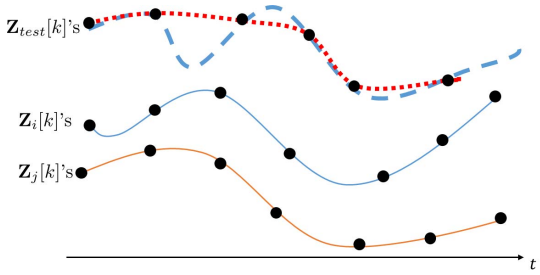


Fig. 6. Illustration of variability in event instances.

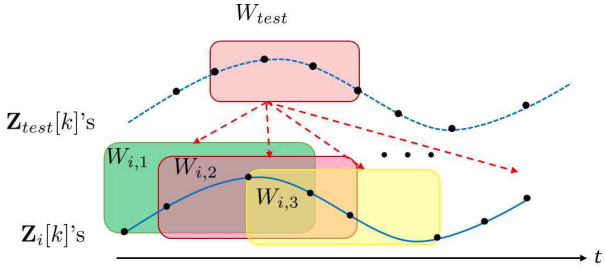


Fig. 7. Illustration of monitoring with partial information by sliding window.

instances and difficulties in locating either the start or the end point of event occurrence in the incoming testing stream.

1) Challenges in Real-Time Monitoring: Real-time indoor monitoring faces a lot of challenges. On one hand, considering the low latency requirement for a practical indoor monitoring system, it is hard to locate either the start or the end point of an event occurrence in the incoming infinite testing CSI stream. The real-time indoor monitoring system should be able to detect the trained event promptly, even before the event stops. Furthermore, in the course of daily monitoring, the event occurrence may be halted due to unknown reason. On the other hand, because the duration of different events may vary, the length of training feature series is different. In order to have a fair similarity comparison between the same testing series and different training series with varied lengths, the training feature sequences need to be trimmed to the same length. As one of the typical trimming methods, downsampling over training sequences fails to meet the requirement because of information loss. Unlike downsampling, the original information can be preserved and the problem of different training feature lengths will be resolved by dividing the training series of varied length into several equal-length subsequences. As only a part of training information is contained in each subsequence, the proposed system is required to be able to perform high accuracy classification over partial training information.

On the other hand, in practice, the manner of how an event occurs and evolves will be different when performed by different individuals, resulting in a testing feature series different from the training one. Asynchronized sampling during the event occurrence that is continuous in time also leads to an altered testing feature. The proposed system should be capable of handling variabilities in event instances as discussed above.

An example that illustrates the variabilities in event instances is shown in Fig. 6. Incoming testing series $\mathbf{Z}_{\text{test}}[k]$'s is denoted by dashed curves where sampled points are marked by black dots. The bottom two curves represent the fixed-length training series $\mathbf{Z}_i[k]$'s and $\mathbf{Z}_j[k]$'s. By comparing the dashed curve with both solid curves, it is observed that $\mathbf{Z}_{\text{test}}[k]$'s contains full information of the curve denoted by $\mathbf{Z}_i[k]$'s. However, because of sampling problem, the sampled version of $\mathbf{Z}_{\text{test}}[k]$'s exhibits a pattern in the red dotted curve, which is similar to $\mathbf{Z}_j[k]$'s.

2) Monitoring With Partial Training Information: To address the first challenge of partial information monitoring, in this paper, a sliding window with length T_{win} is applied over the incoming testing stream. In practice, the window length T_{win} is selected to be 2 or 3 s, considering the fact that typically an indoor event last for at least 3 s.

As demonstrated in Fig. 7, the incoming testing stream $\mathbf{Z}_{\text{test}}[k]$'s passes through a sliding window and the newest sample with its $T_{\text{win}}/T_s - 1$ preceding samples from the current testing window W_{test}. Similar to that, the training series $\mathbf{Z}_i[l]$'s for every event is also partitioned into several shortened training window, denoted as $W_{i,l}$'s with $W_{i,l}$ being the lth training window for training series $\mathbf{Z}_i[l]$'s. Then the similarity comparison is made between testing window W_{test} and all training windows $W_{i,l}$'s, for all i's and l's.

By applying a sliding window over the income testing stream, the proposed system is able to promptly report the current indoor state. In practice, the length of the stride for sliding window, i.e., the number of antecedent feature samples to be included in W_{test} is set to be $1/T_s$ and the overlap between consecutive training window $W_{i,1}$ and $W_{i,2}$ is set to be $1/2T_s$, for the purpose of avoiding misdetection and unnecessary calculation complexity.

3) Similarity Comparison With Event Variability: In this part, we will present how the proposed system detects current indoor state based on the information in W_{test}. The proposed system adopts a two-stage detection algorithm: 1) a dynamics detector works first over W_{test} to see if the environment is static or dynamic and 2) an event detector then works to determine which trained event occurs if the motion detector reports dynamic.

The dynamics detector measures and tracks the variations within W_{test} by

$$\beta_{\text{test}} = \sum_{k=1}^{|W_{\text{test}}|} \left\| \mathbf{Z}_{W,\text{test}}[k] - \mathbf{Z}_{W,\text{test}}[1] \right\|^2 \tag{13}$$

where $\mathbf{Z}_{W,\text{test}}[k]$ denotes the kth sample in W_{test}, $|W_{\text{test}}|$ represents the total number of samples in W_{test}, and β_{test} is the in-window dynamic metric. When $\beta_{\text{test}} \geq \gamma_{\text{dynamic}}$, the proposed dynamics detector considers current indoor environment to be dynamic and the event detector will respond and work.

Once the system detects dynamics in the environment, it will determine the current indoor event by comparing W_{test} with all training templates $W_{i,l}$, $\forall i \& l$. Taking into consideration of possible variabilities in event instances, in this paper, we adopt DTW to measure the similarity between testing and training windows. As proposed in [39] and [40], DTW adopts dynamic programming to obtain minimum distance alignment between two time series. The DTW distance is indeed the Euclidean

distance between two time series, calculated along the optimal warping path and under the boundary conditions as well as global constraints.

In the proposed algorithm, given two sequences of feature series $\mathbf{Z}_1[l]$'s and $\mathbf{Z}_2[l]$'s with equal length L, the DTW optimal cost c is defined as the normalized distance of a warping path, i.e.,

$$c(\mathbf{Z}_1, \mathbf{Z}_2) = \frac{1}{|P^*|} \sum_{w=1}^{|P^*|} \left\| \mathbf{Z}_1\left[k^*_{1,w}\right] - \mathbf{Z}_2\left[k^*_{2,w}\right] \right\|^2 \qquad (14)$$

where P^* denotes the optimal warping path with length $|P^*|$, and $k^*_{1,w}$ and $k^*_{2,w}$ are the indexes of $\mathbf{Z}_1[k]$'s and $\mathbf{Z}_2[k]$'s at the wth point on the path P^*.

For all possible warping paths P with $(l_{1,w}, l_{2,w})$'s, P^* is the optimal one in that

$$\frac{1}{|P^*|} \sum_{w=1}^{|P^*|} \left\| \mathbf{Z}_1\left[k^*_{1,w}\right] - \mathbf{Z}_2\left[k^*_{2,w}\right] \right\|^2 \leq \frac{1}{|P|} \sum_{w=1}^{|P|} \left\| \mathbf{Z}_1\left[k_{1,w}\right] - \mathbf{Z}_2\left[k_{2,w}\right] \right\|^2 \quad \forall P. \qquad (15)$$

With a warping step-size $\Delta > 1$, i.e., the allowable largest stepsize for path advancing, the DTW algorithm is able to overcome issues of missing feature samples introduced by variabilities in event instances and WiFi traffic collision. In addition, in the proposed algorithm, the Sakoe-Chiba Band [39] is adopted which reduces the number of searchable indexes and thus the proposed algorithm benefits from a quick and low-complexity computation of DTW.

After that, the final distance between W_{test} and all training templates $\mathbf{Z}_i$'s is defined based on (14) as

$$\hat{c}(W_{\text{test}}, \mathbf{Z}_i) = \min_{W_{i,l} \subseteq \mathbf{Z}_i's} c\left(W_{\text{test}}, W_{i,l}\right). \qquad (16)$$

With the help of DTW, a simple kNN classifier is sufficient to classify testing window W_{test}. The decision rule is as follows:

$$D_{\text{test}}(W_{\text{test}}) = \begin{cases} \arg\min_{S_i \in S} \ \hat{c}(W_{\text{test}}, \mathbf{Z}_i) \\ \text{if } \min_{S_i \in S} \ \hat{c}(W_{\text{test}}, \mathbf{Z}_i) \leq \gamma_{\text{event}} \\ \text{Unknown, otherwise} \end{cases} \qquad (17)$$

where γ_{event} is an empirical threshold and $D_{\text{test}}(W_{\text{test}}) =$ Unknown indicates the occurrence of an unrecognized indoor event or indoor state.

C. Algorithms for Unsupervised Retraining

Another big challenge for real-time wireless indoor monitoring is unpredictable and inevitable changes in the indoor propagation environment. Due to normal human activities inside the monitored areas and channel fading, the estimated multipath CSI keeps changing along time and it may result in a mismatch between testing and training feature series. It is crucial to design a real-time indoor monitoring system that can adapt itself to environment changes and maintain its performance in long-term. In this section, we propose an unsupervised automatic retraining algorithm that keeps updating the training database of the proposed system on the fly. The proposed algorithm contains two part: 1) retraining for static environment and 2) retraining for dynamic events. The details are as follows.

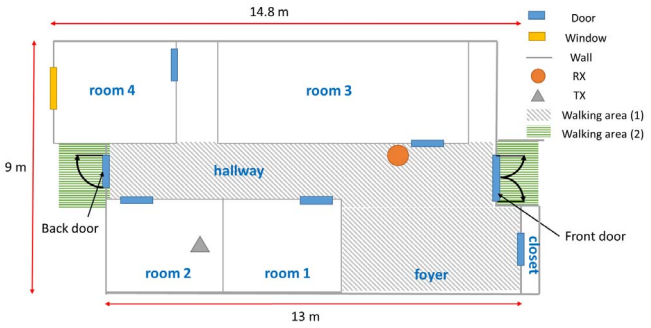


Fig. 8. Experimental setup for setting 1.

Algorithm 2 Unsupervised Retraining for Dynamic Events

Input: Incoming testing stream $\mathbf{Z}_{W,test}[k]$'s as in W_{test}, D_{test} for W_{test}.
Output: Candidate training series $\widehat{W}$.

Initiate $W_{\text{buffer}} \leftarrow$ empty matrix of $p_c \times 0$.
$W_{\text{buffer}} \leftarrow \mathbf{Z}_{W,test}[k]$, if $\hat{c}(W_{test}, \mathbf{Z}_{D_{test}}) \leq \gamma_{un}$.
while TRUE do
  $D_{prev} \leftarrow D_{test}$.
  $W_{test} \leftarrow$ new testing features $\mathbf{Z}_{W,test}[k]$'s.
  Obtain D_{test} for current W_{test}.
  if W_{buffer} is empty **then**
    if $\hat{c}(W_{test}, \mathbf{Z}_{D_{test}}) \leq \gamma_{un}$ **then**
      $W_{\text{buffer}} \leftarrow \mathbf{Z}_{W,test}[k]$.
    end if
  else
    if $D_{prev} == D_{test}$ and $D_{test} \in S$ and $\hat{c}(W_{test}, \mathbf{Z}_{D_{test}}) \leq \gamma_{un}$ **then**
      $W_{\text{buffer}} \leftarrow$ concatenate W_{buffer} and $\mathbf{Z}_{W,test}[k]$'s with repeated testing feature vectors discarded.
    else
      $W_{\text{buffer}} \leftarrow$ feature extraction over Gaussian filtered W_{buffer}.
      if $|W_{\text{buffer}}| \geq |W_{test}|$ **then**
        $\widehat{W} \leftarrow W_{\text{buffer}}$ for event D_{test} and is put into training database.
      end if
      $W_{\text{buffer}} \leftarrow$ empty matrix of $p_c \times 0$.
    end if
  end if
end while

1) *Retraining for Static Environment:* Feature sequence from the static environment can be easily and reliably labeled through the proposed dynamic detector. Hence, in the proposed system, the training series $\mathbf{Z}_i[l]$'s that represent a static environment is periodically updated with the testing series W_{test} which is classified as a static state by the proposed system. In order to guarantee the robustness of retraining for the static state, a boundary condition is set for the candidate feature series W_{test} to be qualified in that the decision outputs before and after W_{test} must be static consistently for a certain period (e.g., 1 h).
2) *Retraining for Dynamic Events:* The unsupervised retraining procedure for trained dynamic events is much

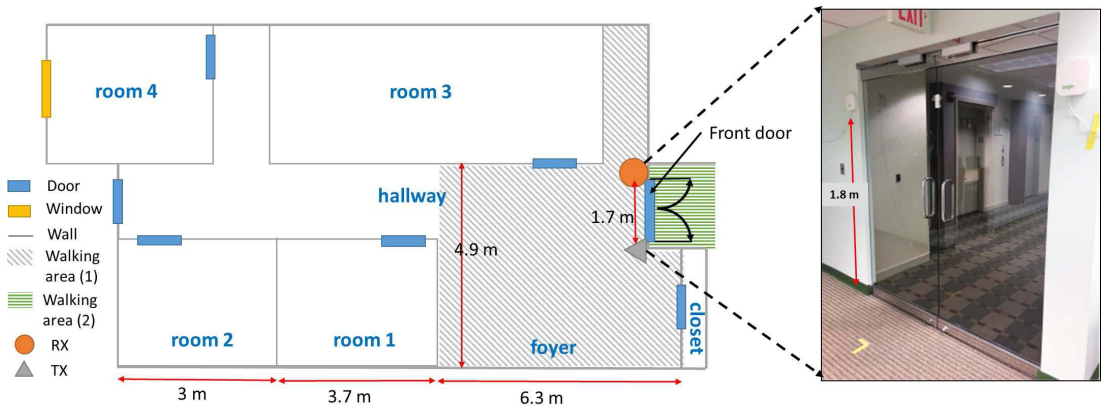


Fig. 9. Experimental setup for setting 2.

more complicated. The criterion to select a qualified testing series as a new training series for an event can either be too loose which may introduce more false alarms to the system or be too strict which may reject all possible candidates and make the proposed system incapable of self-adapting to environmental changes. In the proposed system, the retraining for dynamic events works under the following protocols as listed in Algorithm 2.

During the monitoring phase, whenever the system detects the current indoor states to be a trained events and the distance score $\hat{c}(W_{\text{test}}, \mathbf{Z}_{D_{\text{test}}}) \leq \gamma_{\text{un}}$, the corresponding testing series W_{test} will be stored temporarily in a buffer W_{buffer} and the system decision is put into D_{prev}. Subsequent W_{test}'s will be concatenated into the same buffer W_{buffer} with repeated feature samples being discarded, if their decision D_{test} is as same as D_{prev} and the distance score satisfies $\hat{c}(W_{\text{test}}, \mathbf{Z}_{D_{\text{test}}}) \leq \gamma_{\text{un}}$. When the current D_{test} is different from D_{prev} or $\hat{c}(W_{\text{test}}, \mathbf{Z}_{D_{\text{test}}}) \geq \gamma_{\text{un}}$, one will apply the proposed transition extraction algorithm over the stored testing features in W_{buffer}, and put the extracted feature sequence into the database as a new training sequence $\widehat{W}$ for event D_{test} only if the minimum length criterion is satisfied. Lastly, the buffer W_{buffer} is reset and ready for the next coming testing series. In Algorithm 2, γ_{un} is an empirical threshold, chosen to satisfy $\gamma_{\text{un}} \leq \gamma_{\text{event}}$ such that the confidence of testing samples in W_{buffer} belonging to event D_{test} is guaranteed. Moreover, as the system works along time, the number of training series for an event grows. Hence, the proposed system keeps forgetting the oldest training series for each event obtained from Algorithm 2 when the total number of training series for that event exceeds a predefined capacity.

IV. Experimental Results

To evaluate the performance of the proposed algorithms, extensive experiments have been conducted to protect a multiroom office from intrusion in various settings. In this paper, we use the door open/closed event monitoring as an example to illustrate the performance of the proposed algorithms. The proposed real-time indoor monitoring system can be extended to other applications. Contact sensors, as well as a video recording system, are used to provide the ground truth in the experiments.

A. Experimental Setup

A prototype of the proposed indoor monitoring system is implemented using a pair of commercial WiFi devices, which performs 2×3 MIMO transmission with the carrier frequency being 5.8 GHz and under a 40-MHz bandwidth.

The experiments are carried out in the offices at the 10th floor in a commercial building of 16 floors in total. The experimental offices are surrounded by multiple offices and four elevators. The experimental setting is shown in the floorplans in Figs. 8 and 9 with location of the TX and RX marked. In setting 1 the system is aimed to monitor both front door and back door opening events, while the system only monitors the front door in setting 2. The window length $|W_{\text{test}}|$ is set to be 3 s.

We divide the experiment into two parts. First, the robustness of the proposed system is evaluated under the variabilities in event instances (Section IV-B), the existence of outdoor activities (Section IV-C), and the existence of indoor activities (Section IV-D). Then, the long-term performance of the proposed system is studied in an experiment lasting for 32 days.

B. Robustness to Event Variability

In this part, we deploy the system in setting 1 (as shown in Fig. 8) and in the training phase, one trainer performs door opening event from door close to door 90-degree open, at both the front door and the back door once with a normal speed. Then, in the testing phase, the tester, other than the trainer, intentionally introduces variabilities in event instances, by opening the door at the same training speed, twice of the training speed, and half of the training speed. Moreover, the tester also performs the door opening event at the same training speed but pauses at a 45-degree. The system output is shown in Fig. 10, where the x-axis is the time index in seconds and the y-axis is the event name.

From Fig. 10, we have the following observations.

1) When the speed of door opening is slow, the proposed system sometime may consider the current environment as static.
2) By leveraging DTW, the proposed system can handle the difference in training and testing speed.

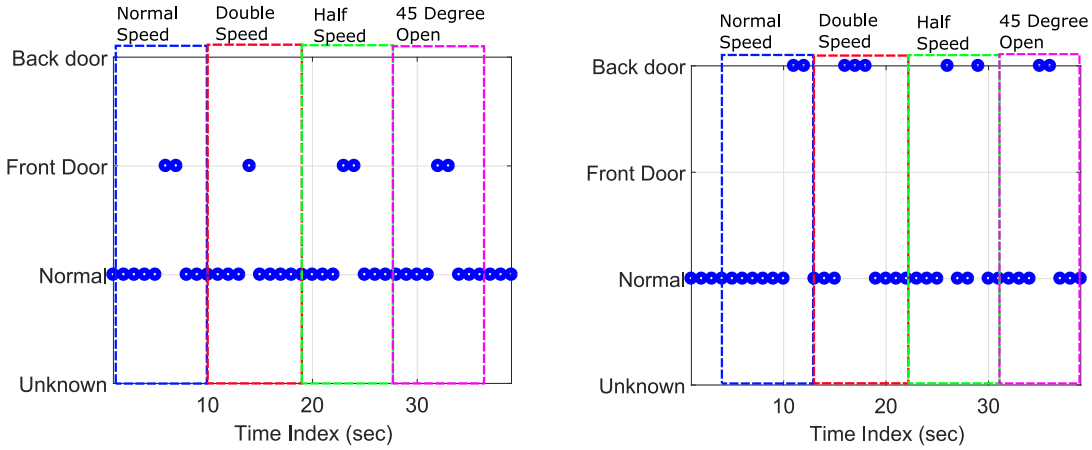


Fig. 10. Experimental results on the robustness to variabilities in event instances in setting 1. (a) System output for front door opening test. (b) System output for back door opening test.

3) With the help of the sliding window, the proposed system can detect indoor state with partial training information. Hence, even when the door opening process is paused at half open which is the middle point of the training series, the proposed system can still reliably and promptly detect the occurrence of the event.

In general, the proposed system is robust to variabilities in event instances.

C. Robustness to Outside Activities

In this part, we conduct experiments to evaluate how the system responds to outside activities. The experimental setting is the same as the one in Figs. 8 and 9, where the tester is walking randomly in the horizontally shaded area [walking area (2)], close to but outside the target event. The experiment simulates a postman scenario.

With zero false alarm generated during the 5-min postman experiment under both settings, the proposed real-time indoor monitoring system is insusceptible to activities outside the monitored area.

D. Robustness to Inside Activities

Because the amount of multipath energy leaks to the outside of the monitored area is limited, the proposed real-time indoor monitoring system is robust to outside activities. However, as most of the multipaths concentrate inside the monitored area and especially around the TX/RX device, inside activities perturb the measured multipath CSI and may introduce a lot of false alarms, posing a great challenge to the wireless monitoring. In this part, we conduct three sets of experiments to test the robustness of the proposed system to inside activities.

1) In Scenario 1, whose setting is shown in Fig. 8, one tester is asked to walk randomly in the foyer and along the hallway, as shaded by the diagonal lines. The system is trained only with the front and the back door opening events. During the 1-min inside walking test, no false alarm is triggered and the system works correctly along time.

TABLE II
LONG-TERM TEST RESULTS

Total Duration of Monitoring (seconds)	1,968,420
Total Number of Event Occurrence	2085
Number of Detected Event Occurrence	2078
Detection Rate	99.66%
Number of False Alarm (seconds)	17
False Alarm Rate	$8.64e-6$

2) We then deploy the system under setting 2 and train it to monitor the front door opening event. We start with the case when one and two testers are asked to walk randomly in the diagonally shaded area as shown in Fig. 9, which is very close to the trained event and the TX/RX. In the 6-min test, the system output is correct no matter how many people are walking inside the monitored area.
3) In addition to the test of human walking in the shaded area, we also perform a different door opening test to investigate whether a similar but untrained event will introduce false alarms to the proposed system. In this scenario, one tester is asked to open and close doors of room 3 and the closet four times, respectively. The door locations are marked in Fig. 9 and both doors are near the targeted front door. Nevertheless, according to the system output during the test, no false alarm is reported and the proposed system is robust in distinguishing between similar events even without full prior knowledge.

Based on the experimental experiments in this part, we can conclude that the proposed real-time monitoring system guarantees its robustness to various indoor activities.

E. Long-Term Performance

To evaluate the long-term performance and the unsupervised retraining algorithms of the proposed system, the prototype is deployed in setting 2 for 32 days without human intervention. The system is aimed to protect the front door and only

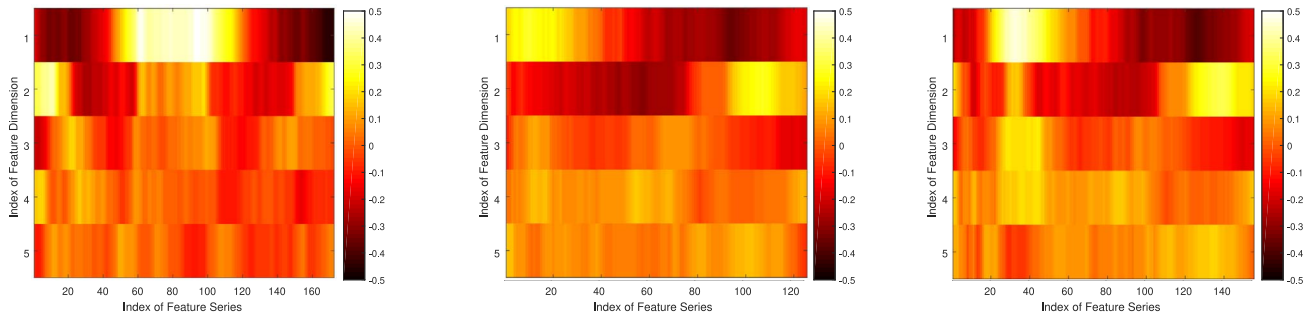


Fig. 11. Examples of feature series generated by Algorithm 2. (a) Original feature series. (b) Example 1. (c) Example 2.

gets trained once on the first day. The front door is the major entrance for the entire office area which is occupied by over 12 people. Moreover, during the testing, normal workday activities happen inside the testing area and the furniture inside the foyer and room 3 gets moved from time to time. The result is summarized in Table II.

According to the results listed in Table II, the proposed system succeeds in maintaining a high accuracy of trained event detection with the help of the proposed automatic unsupervised training algorithm. Specifically, during the 547-h of real-time monitoring test, the proposed system has a detection rate of 99.66% while the false alarm rate is only $8.64e{-}6$, i.e., only has 17 s of false alarms out of 547 h. To the best of our knowledge, this paper is the first real-time indoor monitoring system performing fine-grained event detection that has been tested in a busy office environment for over one month without human intervention.

During the 32-day experiments, the front door opening event has been updated for 475 times in an unsupervised manner, while the latest 19 and the original one are kept in the database. Examples of new training feature series obtained from unsupervised retraining during the long-term test are shown in Fig. 11. Although the same trend exhibits in those three sequences, the new feature series in either Fig. 11(b) or (c) is slightly different from the original training series in Fig. 11(a) captured at the initialization.

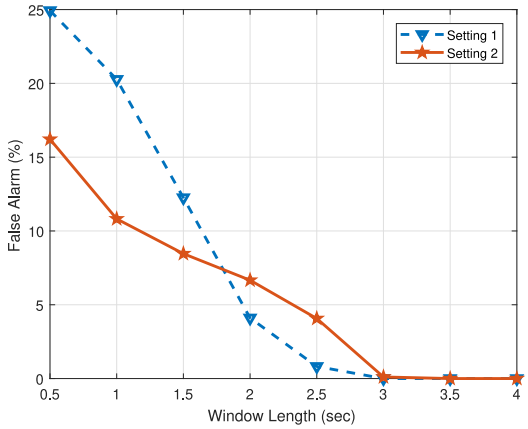


Fig. 12. Experimental results on the impact of sliding window length.

TABLE III
LONG-TERM TEST RESULT WITHOUT UNSUPERVISED RETRAINING FOR DYNAMIC EVENTS

Total Duration of Monitoring (seconds)	1,968,420
Total Number of Event Occurrence	2085
Number of Detected Event Occurrence	1844
Detection Rate	88.44%
Number of False Alarm (seconds)	2
False Alarm Rate	$5.08e-7$

V. DISCUSSION

In this section, we will study the impact of the length of the sliding window W_{test} and the proposed unsupervised retraining algorithm. We will also demonstrate how the proposed system can be utilized for future smart home or smart office applications.

A. Impact of Sliding Window Length

We now revisit the experimental results in Section IV-D and evaluate with different lengths of W_{test} to study how different lengths of W_{test} will affect system performance. The false alarm rate is studied based on different window length $|W_{\text{test}}|$, and the result is plotted in Fig. 12 for setting 1 and setting 2.

According to the result in Fig. 12, for both settings, the false alarm rate increases significantly as the window length $|W_{\text{test}}|$ decreases below 3 s. That is because the distinctive information of events is embedded not only in each instantaneous CSI sample but also in the transition information of the CSI time sequence. As we increase the window length $|W_{\text{test}}|$, more environmental information in the CSI is included for event recognition. Although two events may share resembling individual CSI samples, the associated CSI time sequence should be differentiable. Hence, as more information is included as the representative pattern, higher detection accuracy can be achieved and fewer false alarms will be triggered by nearby human interference. On the other hand, with a larger $|W_{\text{test}}|$, the system latency increases which is undesirable for a real-time system. Hence, in order to guarantee the performance of the proposed system, the window length $|W_{\text{test}}|$ should be carefully selected.

B. Impact of Unsupervised Retraining

In this part, the long-term system performance with and without the proposed unsupervised retraining algorithms is

TABLE IV
COMPARISON ON RECENT CSI-BASED INDOOR EVENT (OBJECT STATE) DETECTION SYSTEMS FOR SURVEILLANCE APPLICATIONS

Methodology	Accuracy	Set-Up Effort	Limitations
Time-reversal based similarity comparison between individual CSI [35]	96.92%	Low	Characterizes each event as a static state Lack of long-term study Lack of evaluation on real-time detection Not integrated with commercial WiFi devices
Statistics of intra-class CSI [36]	95.45%	Medium	Characterizes each event as a single static state
Neural network and hidden Markov model based classification on CSI [20]	85%	High	Thousands of labeled training data required High computation complexity 1000 Hz sounding rate Lack of evaluation on real-time detection
SVM classification over wavelet feature from CSI [26]	94.5%	Medium	100 Hz sounding rate Loss of CSI phase information Lack of long-term study Hundreds of training samples for each event Lack of evaluation on real-time detection

compared and discussed. The result without unsupervised retraining for dynamic events is summarized in Table III.

Comparing Tables II and III, the proposed unsupervised retraining algorithm helps to maintain the detection performance in long-term when the background environment changes along time. Without the proposed unsupervised retraining, the detection rate drops from 99.66% to 88.44%. However, because of the nature of unsupervised labeling, the candidate training series extracted from testing CSI stream may introduce slight false alarms coming from the interference of human activities happening before or after the trained events. Overall, the proposed system equipped with the automatic unsupervised retraining scheme is promising for real-time indoor monitoring applications.

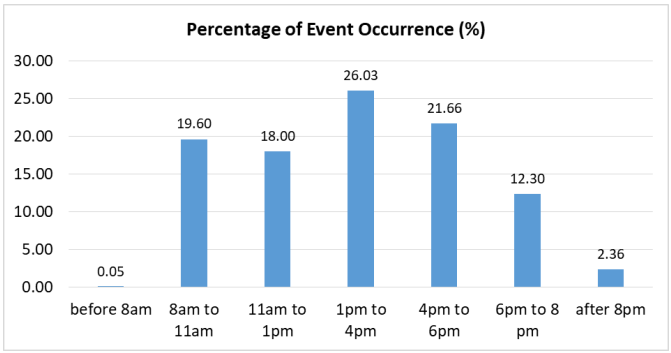


Fig. 13. Smart office analysis.

C. Potential Application for Smart Home/Office

In this part, we discuss how the proposed real-time indoor monitoring system can be deployed in future smart home/office. By performing analytics over the monitoring results generated by the proposed system, we can analyze human behaviors in the monitored area and get the following chart as shown in Fig. 13.

Fig. 13 indicates active and inactive hours of the office area where the long-term experiment is conducted. That measured human behavioral information can incorporate with such applications as smart air conditioner, smart lights, and other devices, to create a smart office environment, which is energy efficient and user-friendly.

D. Detection of Large Number of Events

In this paper, we propose an indoor monitoring system based on the feature extracted from the CSI time series to detect and differentiate between different events in real time. Different kinds of events involve objects moving along distinct trajectories in the physical space. The changes due to the occurrence of an event will be reflected in the CSI as discussed in Section II. The original feature vector in the time series has a dimension of $N_{\mathrm{TX}} \times N_{\mathrm{RX}} \times L$, where N_{TX}, N_{RX}, and L represent the number of transmit antenna, the number of receive antenna and the number of accessible subcarriers in the CSI, respectively. Given the large feature dimension obtained from the prototype mentioned in Section IV-A, the proposed system should be able to detect multiple different events which bring in different patterns to the CSI time series.

However, a critical issue in the indoor monitoring system is that the perturbations or changes an indoor event introduces to the CSI are negligible when it happens far away from the WiFi transmitter-receiver link. This may result in a failure of detecting that event. The impact of the locations of the transmitter and the receiver on the indoor event detection system has been studied in [36]. In order to detect multiple kinds of events ($>>2$) happening in a large indoor area, one solution is to deploy multiple transmitters or receivers distributively. With multiple WiFi receiver-transmitter links available in the space, we not only increase the feature dimension and degrees of freedom that enables to capture and distinguish more fine-grained CSI changes, but also enlarge the coverage of the wireless sensing system.

E. Comparison With Existing Approaches

In Table IV, we compare the existing CSI-based indoor event detection systems for surveillance applications in terms of the methodology, average accuracy, and limitations. Compared with the existing works, the system we propose in this paper characterizes each indoor event as a unique dynamic transition, whose changes are recorded in the CSI time series. To recognize different indoor events, algorithms

are designed to support real-time monitoring with high accuracy and the capability of adapting to practical environmental changes. No device calibration is required and the proposed system can be implemented using commercial WiFi devices. During a one-month evaluation deployed in a real office environment, the proposed system can achieve an accuracy of over 99%, which to our knowledge outperforms the existing approaches.

Moreover, we also compare systems listed in Table IV, in terms of the effort required for making the system detecting accurately when being deployed in a new environment, also known as, the set-up effort. The system proposed in [20] requires the highest effort due to the training for the deep learning network. The set-up effort for deploying the systems proposed in [26] and [36] in a new environment is medium, because hundreds of training samples are required to generalize features and build the classifier. On the other hand, the system proposed in this paper is of low set-up effort. When being deployed in a new environment, only one training sequence for each event is needed and the system is able to collect more training sequences during the real-time monitoring phase in an unsupervised manner.

VI. Conclusion

In this paper, we have proposed a real-time indoor event monitoring system that utilizes CSI time series to differentiate between indoor states. We have designed a feature extraction algorithm to extract and refine a low-dimensional feature from the measured CFR sequences. To address practical issues of real-time monitoring, including variabilities in event instances and unknown start and end point in the incoming testing stream, we have proposed a sliding window-based classification algorithm with DTW measuring the similarity between training and testing features. We also have developed an automatic unsupervised retraining algorithm to improve the system performance in long-time monitoring. Extensive experiments were conducted and the proposed system achieved a detection rate of 99.66% with a false alarm rate of only $8.64e{-}6$ during a 547-h of monitoring test. Experimental results further demonstrate the potential of the proposed system in future real-time indoor monitoring applications.

References

[1] B. Wang, Q. Xu, C. Chen, F. Zhang, and K. J. R. Liu, "The promise of radio analytics: A future paradigm of wireless positioning, tracking, and sensing," *IEEE Signal Process. Mag.*, vol. 35, no. 3, pp. 59–80, May 2018.

[2] M. Spadacini, S. Savazzi, M. Nicoli, and S. Nicoli, "Wireless networks for smart surveillance: Technologies, protocol design and experiments," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, 2012, pp. 214–219.

[3] G. Fortino and P. Trunfio, *Internet of Things Based on Smart Objects: Technology, Middleware and Applications*, 1st ed. Cham, Switzerland: Springer, 2014.

[4] D. Zhang, J. Ma, Q. Chen, and L. M. Ni, "An RF-based system for tracking transceiver-free objects," in *Proc. 5th Annu. IEEE Int. Conf. Pervasive Comput. Commun.*, Mar. 2007, pp. 135–144.

[5] C. Han, K. Wu, Y. Wang, and L. Ni, "WiFall: Device-free fall detection by wireless networks," in *Proc. Int. Conf. Comput. Commun.*, Apr. 2014, pp. 271–279.

[6] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2015, pp. 1472–1480.

[7] Y. Gu, F. Ren, and J. Li, "PAWS: Passive human activity recognition based on WiFi ambient signals," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 796–805, Oct. 2016.

[8] K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, "Enhanced fingerprinting and trajectory prediction for IoT localization in smart buildings," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 3, pp. 1294–1307, Jul. 2016.

[9] L. Yang, W. Li, M. Ghandehari, and G. Fortino, "People-centric cognitive Internet of Things for the quantitative analysis of environmental exposure," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2353–2366, Aug. 2018.

[10] C. Wu *et al.*, "Non-invasive detection of moving and stationary human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2329–2342, Nov. 2015.

[11] Y. Zeng, P. H. Pathak, C. Xu, and P. Mohapatra, "Your AP knows how you move: Fine-grained device motion recognition through WiFi," in *Proc. 1st ACM Workshop Hot Topics Wireless*, 2014, pp. 49–54.

[12] Y. Wang *et al.*, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proc. 20th ACM Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 617–628.

[13] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial WiFi devices," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1118–1131, May 2017.

[14] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using WiFi signals," in *Proc. 21st ACM Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 90–102.

[15] S. Tan and J. Yang, "WiFinger: Leveraging commodity WiFi for fine-grained finger gesture recognition," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2016, pp. 201–210.

[16] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, "PADS: Passive detection of moving targets with dynamic speed using PHY layer information," in *Proc. 20th IEEE Int. Conf. Parallel Distrib. Syst.*, Dec. 2014, pp. 1–8.

[17] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. Ni, "FIMD: Fine-grained device-free motion detection," in *Proc. 18th Int. Conf. Parallel Distrib. Syst.*, Dec. 2012, pp. 229–235.

[18] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. 21st ACM Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 65–76.

[19] H. Zhu, F. Xiao, L. Sun, R. Wang, and P. Yang, "R-TTWD: Robust device-free through-the-wall detection of moving human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1090–1103, May 2017.

[20] K. Ohara, T. Maekawa, and Y. Matsushita, "Detecting state changes of indoor everyday objects using Wi-Fi channel state information," in *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, Sep. 2017, p. 88.

[21] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th ACM Annu. Int. Conf. Mobile Comput. Netw.*, 2013, pp. 27–38.

[22] B. Tan *et al.*, "Wi-Fi based passive human motion sensing for in-home healthcare applications," in *Proc. IEEE 2nd World Forum Internet Things (WF-IoT)*, 2015, pp. 609–614.

[23] K. Qian *et al.*, "Inferring motion direction using commodity WiFi for interactive exergames," in *Proc. ACM Conf. Human Factors Comput. Syst. (CHI)*, 2017, pp. 1961–1972.

[24] Y.-X. Zhao, N.-X. Du, Z.-Y. Fang, and H.-Y. Sun, "Design and implementation of frequency domain feature extraction module for human motion recognition," in *Proc. Int. Conf. Wireless Commun. Netw. Appl.*, 2017, pp. 51–56.

[25] K. Qian *et al.*, "Enabling contactless detection of moving humans with dynamic speeds using CSI," *ACM Trans. Embedded Comput. Syst.*, vol. 17, no. 2, p. 52, Apr. 2018.

[26] M. A. A. Al-Qaness, F. Li, X. Ma, and G. Liu, "Device-free home intruder detection and alarm system using WiFi channel state information," *Int. J. Future Comput. Commun.*, vol. 5, no. 4, p. 180, 2016.

[27] K. Qian, C. Wu, Z. Yang, C. Yang, and Y. Liu, "Decimeter level passive tracking with WiFi," in *Proc. 3rd Workshop Hot Topics Wireless*, 2016, pp. 44–48.

[28] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Trans. Graph.*, vol. 34, no. 6, p. 219, Oct. 2015.

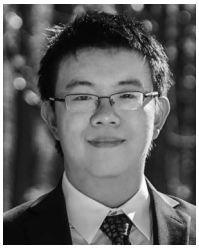
[29] M. Zhao *et al.*, "Through-wall human pose estimation using radio signals," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 7356–7365.

[30] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via RF body reflections," in *Proc. 12th USENIX Symp. Netw. Syst. Design Implement.*, May 2015, pp. 279–292.
[31] G. K. Nanani and M. Kantipudi, "A study of WiFi based system for moving object detection through the wall," *Int. J. Comput. Appl.*, vol. 79, no. 7, pp. 15–18, 2013.
[32] Y. Zhu, B. Y. Zhao, and H. Zheng, "Reusing 60 GHz radios for mobile radar imaging," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 103–116.
[33] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D tracking via body radio reflections," in *Proc. 11th USENIX Symp. Netw. Syst. Design Implement.*, Apr. 2014, pp. 317–329.
[34] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. 33rd Annu. ACM Conf. Human Factors Comput. Syst.*, 2015, pp. 837–846.
[35] Q. Xu, Y. Chen, B. Wang, and K. J. R. Liu, "TRIEDS: Wireless events detection through the wall," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 723–735, Jun. 2017.
[36] Q. Xu, Z. Safar, Y. Han, B. Wang, and K. J. R. Liu, "Statistical learning over time-reversal space for indoor monitoring system," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 970–983, Apr. 2018.
[37] Q. Xu, C. Jiang, Y. Han, B. Wang, and K. J. R. Liu, "Waveforming: An overview with beamforming," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 132–149, 1st Quart., 2018.
[38] Q. Xu, Y. Chen, B. Wang, and K. J. R. Liu, "Radio biometrics: Human recognition through a wall," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 5, pp. 1141–1155, May 2017.
[39] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-26, no. 1, pp. 43–49, Feb. 1978.
[40] M. Müller, *Information Retrieval for Music and Motion*. Heidelberg, Germany: Springer-Verlag, 2007.

Qinyi Xu (S'15) received the B.S. degree (Highest Hons.) in information engineering from Southeast University, Nanjing, China, in 2013, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland at College Park, College Park, MD, USA, in 2016 and 2018, respectively.

Since 2018, she has been with Origin Wireless Inc., Greenbelt, MD, USA, where she is a Principal Scientist. She is also with the Department of Electrical and Computer Engineering, University of Maryland at College Park. She was an Exchange Student with the KTH Royal Institute of Technology, Stockholm, Sweden, from 2012 to 2013, under the National Sponsorship of China. Her current research interests include signal processing, machine learning, wireless sensing, and wireless communications.

Dr. Xu was a recipient of the Clark School Distinguished Graduate Fellowships from the University of Maryland at College Park and the Graduate with Honor Award from Southeast University in 2013.



Yi Han received the B.S. degree (Highest Hons.) in electrical engineering from Zhejiang University, Hangzhou, China, in 2011, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD, USA, in 2016.

He is currently the Wireless Architect of Origin Wireless, Greenbelt, MD, USA. His current research interests include wireless communication and signal processing.

Dr. Han was a recipient of the Class A Scholarship from Chu Kochen Honors College, Zhejiang University in 2008 and the Best Student Paper Award of IEEE ICASSP in 2016.

Beibei Wang (SM'15) received the B.S. degree (Highest Hons.) in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2004, and the Ph.D. degree in electrical engineering from the University of Maryland at College Park, College Park, MD, USA, in 2009.

She was with the University of Maryland at College Park, as a Research Associate, from 2009 to 2010, and with Qualcomm Research and Development, San Diego, CA, USA, from 2010 to 2014. Since 2015, she has been with Origin Wireless Inc., Greenbelt, MD, USA, where she is currently a Chief Scientist. She co-authored *Cognitive Radio Networking and Security: A Game-Theoretic View* (Cambridge Univ. Press, 2010). Her current research interests include wireless sensing, positioning, machine learning, and communications and signal processing.

Dr. Wang was a recipient of the Graduate School Fellowship, the Future Faculty Fellowship, and the Dean's Doctoral Research Award from the University of Maryland at College Park, and the Overview Paper Award from the IEEE Signal Processing Society in 2015.

Min Wu (S'95–M'01–SM'06–F'11) received the B.E. degree (Highest Hons.) in electrical engineering and the B.A. degree (Highest Hons.) in economics from Tsinghua University, Beijing, China, in 1996, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2001.

Since 2001, she has been with the University of Maryland at College Park, College Park, MD, USA, where she is currently a Professor and the University Distinguished Scholar–Teacher. She leads the Media and Security Team, University of Maryland at College Park. Her current research interests include information security and forensics and multimedia signal processing.

Dr. Wu was a recipient of the NSF CAREER Award in 2002, the TR100 Young Innovator Award from *MIT Technology Review Magazine* in 2004, the ONR Young Investigator Award in 2005, the ComputerWorld 40 Under 40 IT Innovator Award in 2007, the IEEE Mac Van Valkenburg Early Career Teaching Award in 2009, and the IEEE Distinguished Lecturer Recognition in 2015 and 2016. She was a co-recipient of the two Best Paper Awards from the IEEE Signal Processing Society and EURASIP. She has served as the Vice President–Finance of the IEEE Signal Processing Society from 2010 to 2012. She has served as the Chair of the IEEE Technical Committee on Information Forensics and Security from 2012 to 2013. She has also served as the Editor-in-Chief for *IEEE Signal Processing Magazine* from 2015 to 2017. She was elected as an AAAS Fellow for contributions to signal processing, multimedia security, and forensics.

K. J. Ray Liu (F'03) was named a Distinguished Scholar–Teacher with the University of Maryland at College Park, College Park, MD, USA, in 2007, where he is a Christine Kim Eminent Professor of Information Technology and leads the Maryland Signals and Information Group, conducting research encompassing broad areas of information and communications technology with a recent focus on wireless AI.

Dr. Liu was a recipient of the 2016 IEEE Leon K. Kirchmayer Award on graduate teaching and mentoring, the 2014 Society Award from the IEEE Signal Processing Society, the 2009 Technical Achievement Award from the IEEE Signal Processing Society, the Highly Cited Researcher Award by Web of Science, the 2017 CEATEC Grand Prix Award for his invention of the time-reversal machine by Origin Wireless Inc., teaching and research recognitions from the University of Maryland at College Park, including University-Level Invention of the Year Award, the College-Level Poole and Kent Senior Faculty Teaching Award, the Outstanding Faculty Research Award, and the Outstanding Faculty Service Award, all from the A. James Clark School of Engineering, and over a dozen Best Paper Awards. He is the IEEE Vice President of Technical Activities. He was the President of the IEEE Signal Processing Society, where he has served as the Vice President–Publications and Board of Governor, and a member of the IEEE Board of Director as the Division IX Director. He has also served as the Editor-in-Chief for *IEEE Signal Processing Magazine*. He is a Fellow of the AAAS.