WiFi-enabled Occupancy Monitoring in Smart Buildings with a Self-Adaptive Mechanism

Muhammad Salman
Inha University
Incheon, Republic of Korea
salman@inha.edu

Young-Duk Seo†
Inha University
Incheon, Republic of Korea
mysid88@inha.ac.kr

Youngtae Noh†
KENTECH
Naju-si, Republic of Korea
ytnoh@kentech.ac.kr

ABSTRACT

The building energy saving (BES) has been the subject of extensive research for reducing the energy consumption inside the buildings. One of the key solution for energy saving in buildings is to minimize the energy supply to the building areas that are not occupied by the inhabitants. However, this requires effective monitoring of occupants regardless of unpredictable variations in indoor environment, such as variation in the space size, furniture arrangement, nature of occupant's activity (e.g., varied intensities and instances) etc. Currently, various occupancy monitory solutions have been employed in the existing smart buildings, namely PIR sensors, CO_2 sensors, cameras, etc. However, they are costly and sometimes not interoperable to the complex variations in indoor environments. In this paper, we leveraged the fine-grained information of physical layer (*i.e.*, channel state information – CSI) of the commodity WiFi for occupancy detection and developed a self-adoptive method which is interoperable with complex variations in the indoor environment. In indoor contexts of different sized, varied intensities of physical activity, and various instances of activity of daily living (ADL), our testbed evaluation showed an average detection rate of 98.9%, 98.5%, and 98.1%, respectively.

CCS CONCEPTS

• **Information systems** → **Mobile information processing systems**;

KEYWORDS

Channel State Information (CSI), Interoperability, Building Energy Saving (BES), Activity of Daily Living (ADL), multipath effect

ACM Reference Format:
Muhammad Salman, Young-Duk Seo, and Youngtae Noh. 2023. WiFi-enabled Occupancy Monitoring in Smart Buildings with a Self-Adaptive Mechanism. In *The 38th ACM/SIGAPP Symposium on Applied Computing (SAC '23), March 27–March 31, 2023, Tallinn, Estonia.* ACM, Tallinn, Estonia, Article 4, 4 pages. https://doi.org/10.1145/3555776.3577841

1 INTRODUCTION

The world's population is tremendously increasing and it is projected that by the year 2100 the expected population would be 10.4 billion [8]. Such a rapid growth in population would introduce two major concerns for the humans, namely, the shortage of land and run-out energy resources.

To cope with the land demand, the urbanized regions of the world have already adopted living in the buildings to accommodate the huge population density. Usually, such buildings are installed with the heating, ventilation, and air conditioning (HVAC) systems, lighting systems, and systems for supplying hot water, etc., in order to enhance the inhabitants living standard. Unfortunately, building comfort services are inefficiently administered, exacerbating the waste of energy resources (i.e., around 40% of the total energy) [2, 5]. According to a report by MIT, around 30% of the energy consumed by buildings goes wasted [14].

To prevent the energy wastage in modern buildings (a.k.a. smart buildings), variety of systems have been introduced for occupancy monitoring, namely Passive infrared (PIR) sensors, environmental sensors (e.g., CO_2, humidity, and temperature sensor), and cameras. These methods ensure that comfort services are only offered in inhabited areas. However, they require a higher maintenance cost [7], and have lower sensitivity in Non-Line-of-Sight (NLoS) and unreliable performance with a single occupant in large spaces. More recently, the widespread use of WiFi infrastructure in commercial and residential buildings has made the WiFi-based sensing technique more popular for occupancy monitoring [13, 16, 17]. In this paper, we also leverage the WiFi infrastructure and introduce a system supported by any WiFi-enabled IoT device to smartly sense the occupancy of the target zone. However, unlike the existing WiFi-based occupanct sensing technique which underestimates the interoperability to the complex indoor settings, such as variation in size of the target spaces, scatters (e.g., furniture) arrangement, intensities of occupant activities, random instances of activities of daily living (ADL), and temporal variation in the CSI; our system is self-adoptive to all these complex variations in the target environment. The contribution of our paper are three-fold:

- To accurately detect the occupancy in the target zone, we leverage the CSI of the off-the-shelf WiFi. Compared to prevalent solutions, we justify that our occupancy monitoring system is self-adoptive and interoperable to the complex variations in indoor settings.
- To achieve the system's self-adaptability, we use a synergistic model that incorporates the autotuned DBSCAN and control chart technique; wherein the autotuned DBSCAN operates in batch mode and updates the control chart threshold in accordance with the environmental variations.
- To evaluate our system's performance in real-life scenarios, we build a realistic testbed and delivered the result for the complex indoor settings; namely, different room sizes and uncontrolled furniture arrangement, ADL with diverse intensities and instances, and temporal variation of CSI over different periods.

† Corresponding Authors.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SAC'23, March 27 –March 31, 2023, Tallinn, Estonia

© 2023 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9517-5/23/03.
https://doi.org/10.1145/3555776.3577841

2 METHODOLOGY

Channel State Information (CSI) shows how the WiFi signal encounters degradation due to the multipath when it travels from transmitter to the receiver [18]. These details are given at the granularity of the OFDM subcarriers. The presence of occupants in the target region additionally disrupts the WiFi propagation channel, resulting in larger variations in CSI. Exploiting this phenomenon can assist us in determining the movement of occupants in that area. In this section we discuss the self adoptive occupancy monitory process as shown in Figure 1.

2.1 CSI Sanitization

Removal of noisy subcarriers. The number of subcarriers in the received frames varies with the channel width. There are three different CSI subcarrier types present inside the data frames, namely pilor, DC, and data subcarriers. The DC and pilot subcarriers have asymmetric shapes as opposed to data subcarriers. The DC subcarriers create zero bins, while the Pilot subcarriers forms a spiky pattern. As a result, they cause errors in monitoring the occupancy. To alleviate this problem, we discard these subcarriers based on their indices.

2.2 Filtering and Feature extraction

CSI Filtering. There are a lot of unpredictable elements in the indoor environment, like the space size and the availability of scatters etc. These arbitrary elements randomly increase the noises in the CSI subcarriers that eventually cause errors in occupancy detection. To remove the in-band noise (*i.e.*, noise integrated randomly in the CSI subcarriers), the typical filters based on frequency and time domain are not particularly efficient at reducing them [6]. We leverage the Discrete Wavelet Transform (DWT) to effectively remove such noises. To compute the cutoff threshold such that the noise from the in-band CSI data is removed whilst the occupants activity of daily living is preserved, we leverage the following Doppler frequency f_d formula:

$$f_d = \frac{2v}{c} \times f_c. \tag{1}$$

where v is the user velocity, c is the speed of light (*i.e.*, $3 \times 10^8 m/s$), and f_c is the carrier frequency of the commodity WiFi (e.g., 2.4 GHz). Thus the f_d is 24 Hz and 80 Hz for walking (*i.e.*, 1.5 m/s) and running (*i.e.*, 5 m/s) respectively [1]. Using the maximum attainable Doppler frequency (*i.e.*, 80 Hz) as the cutoff threshold for DWT would retain all fluctuation in the CSI data below this frequency and remove noise levels above it.

Feature Extraction. To extract the feature, first we exploit the variation in the magnitude of the CSI's subcarrier. Then we apply Principal Component Analysis (PCA) reduce the dimensionality of the CSI data contained in the sliding window W. The final feature of the CSI data is the first principal component, which is the maximum Eigen value of the Eigen vectors obtained from the correlation matrix of CSI in W.

2.3 Occupancy detection method

The complex variation in the indoor environments have a significant impact on the accuracy of occupancy detection. Herein, we describe our approach for occupancy detection which is self-adaptive to such complex changes in the target scene.

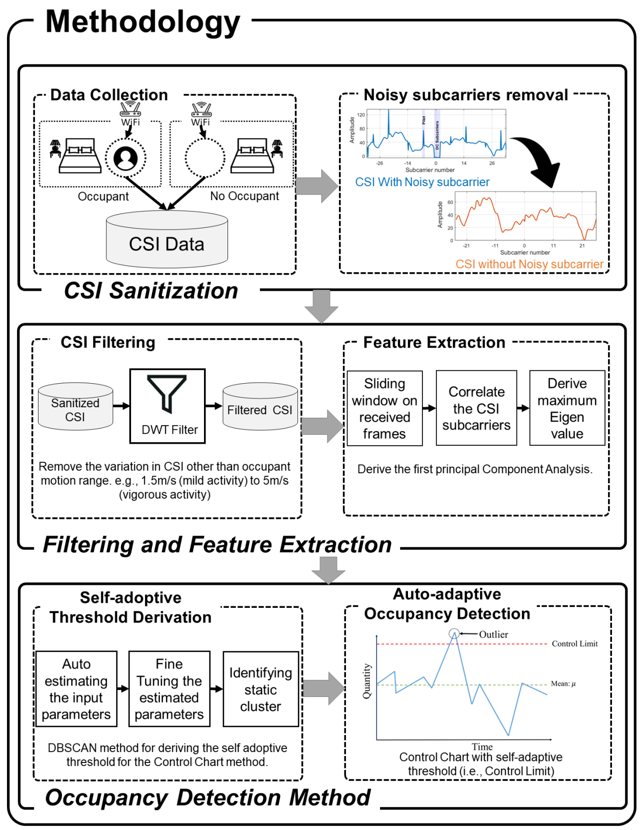


Figure 1: High-level overview of the Occupancy detection methodology.

2.3.1 *Control Chart approach.*

In the control chart technique [15], the time-series quantity is regarded as stable when it stays within the control boundaries (defined by the control limit threshold). In contrast, the quantity is regarded as an outlier when it exceeds the control borders (or control limit). Determining a correct control limit threshold that can reliably identify the presence of occupant in a random indoor-environment is quite challenging. To make the control limit threshold self-adoptive to the changes in random environmental factors, we utilise the DBSCAN method.

2.3.2 *DBSCAN Method.*

The default DBSCAN method [3] requires two input parameters; namely epsilon (ε), and MinPts. The former parameter is the radius of a datapoint, whereas the latter one is the least number of datapoints that must be included in the neighborhood of ε. Because of the random environmental factors, the CSI datapoints substantially change when the ambient environmental changes. Based on these environmental changes the distance between CSI datapoints (*i.e.*, feature value) fluctuates significantly. Owing to the frequent changes in the input parameters of CSI datapoints with the random environmental, the conventional methods for deriving ε (by manually searching the "knee" or first "valley") [3, 4] and MinPts value [4] are infeasible. To overcome this issue, we developed a technique for automatically adopting the MinPts and ε parameters. Our auto-tuned technique involves the following steps.

Auto-estimating the knee point. To determine the CSI data set's ε in arbitrary contexts, we employ the kneedle technique [10]. The fundamental goal of this method is to determine the data set's point which has the maximum curvature. The point (also referred to as knee-point) obtained by this approach is located between datapoints that change rapidly (outliers or dispersed core points) and the datapoints that change slowly (congested core points). In the CSI data set, this is the bifurcation point that separates two clusters (for example, the cluster belonging to non-occupant and occupant cases).

Fine-tuning of the estimated parameters. There are two fundamental flaws in the standard kneedle approach. First, in some indoor environment the multipath effect is very severe, e.g., , a small room. Second, when any of the two activity states (*i.e.*, static and dynamic) has a shortest duration. As a result of these problems, a single cluster is generated for two distinct situations (*i.e.*, non-occupancy and occupancy). To handle these issues, we present a technique for fine-tuning the input parameters (*i.e.*, ε estimated by kneedle and MinPts). First, MinPts is specified to be the same size as the number of dimensions in the dataset as in [4]. Then the ε parameter is iteratively minimized whilst maximizing the MinPts until the two clusters (e.g., non-occupancy data cluster from the occupancy one) are separate. Such ε and MinPts value that causes cluster separation are fine-tuned and can classify the occupancy clusters despite extreme multipath and brief ADL instances.

Determining the non-occupant data cluster. Because fine-tuning enables the formation of at least two clusters, namely non-occupancy and occupancy data clusters. The CSI features belonging to the non-occupancy data are closed to 1, whereas those for occupancy data are less than 1. Furthermore, in contrast to the occupancy case, which is highly variable and dispersed, the cluster produced from the non-occupancy case would have relatively congested datapoints. Based on these observations, the static cluster could be identified.

To sum up, the auto-tuned DBSCAN approach gathers CSI data in batch mode and use them to intelligently and precisely derive thresholds from the CSI's static clusters. This threshold is periodically updated in the control chart method, making it robust to any unanticipated changes in the target environment. As long as, the instantaneous feature value of the CSI remains above this threshold in the control chart method, it as regarded as non-occupancy state, but as this value fall below this threshold, occupant presence is anticipated in the target space.

3 TESTBED SETUP AND EVALUATION

Our testbed configuration is comprised of three device architecture, namely, a transmitter device which sends the traffic, a receiver (access point – AP) which receives the traffic and acknowledge the reception, and a sniffer, which overhears the communication between the transmitter and receiver. The transmitter in our testbed setting was a MacBook Pro laptop, which was pinging 1300 packets/sec [12]. For the receiver, We used a couple of APs in our setting, including ASUS-AC2900, NETGEAR-N600, and TPLINK-AC1750 for the wireless Internet connection. They were all supporting dual band connection. Finally, for the traffic sniffing (*i.e.*, pong traffic from the AP), we used a Raspberry *RPi 3+* (*RPi 3+*) for capturing the CSI. The default WLAN driver of *RPi 3+* does not support the monitoring mode and CSI extraction. To enable these features, we modified its driver using the opensourced GitHub repository of Nexmon [11]. *RPi 3+* collects the CSI (OFDM modulated subcarriers) from the AP using UDP socket 5500 (for listening). We used the occupancy monitoring application both in real-time and offline modes. For the former implementation, we wrote a Python script for occupancy monitoring, and *RPi 3+* was used to execute it via real-time CSI sniffing. Similarly for latter implementation, we transferred the CSI data to the MacBook via USB, and then tested the CSI data set on our occupancy monitoring program.

We evaluated the robustness of our occupancy monitoring system in a variety of random environmental parameters, including (1) variance in indoor settings, (2) occupant activity intensities, (3) ADL instances, and (4) temporal variation in CSI.

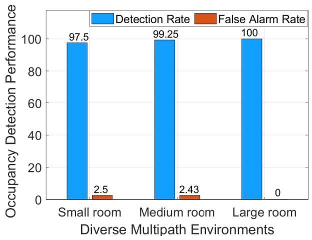
Evaluation of Robustness to the changes in the indoor environment. To assess the robustness of our occupancy monitoring system in random environmental conditions, we consider three rooms, namely small room (with 3.5×3.5 dimension), medium room (with 4×5 dimension), and large room (with 8×8 dimension). Moreover, the environment factors were uncontrolled during the data collection process; more specifically, there was no control over the furniture arrangement, surrounding traffic (which could cause network jitter problem [9]), and movement of people's motion in the vicinity of room. The activity performed by occupant was moderate (*i.e.*, walking in the room at normal pace of 1.4m/s). The data was collected for 30 minutes each time, where the room was empty for initial 10 minutes, and then occupied by the occupant for the latter time (*i.e.*, $10 \sim 30$ minutes).

Figure 2(a) shows the performance of our system in diverse multipath environment. Our occupancy monitoring system performs relatively worse in small room than the other two scenarios. This is due to the shorter inter-wall distance of the small room and the presence of certain scatters (such as furniture) in the target area, which further enhanced the multipath effect. However, the multipath impact reduces as the inter-wall distance increases. As is seen, the detection rate in the large room attained 100% whilst there were no false alarms in the same setting.

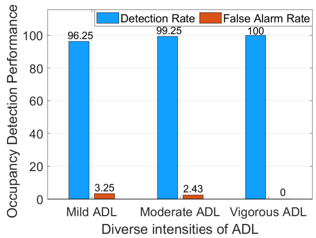
Evaluation of robustness to the intensities of ADL. To assess the robustness of our system with activities of different intensities, we conducted three activities in the medium room for 5 minutes. First, the mild ADL, where the individual was changing a jacket and hanging it on dressing stand and then resting for some time on the chair. Such jacket changing event on average was around 8 times in 5 minutes. Second, moderate ADL, where the individual was consistently walking in the room at pace of 1.5m/s. Finally, vigorous activity, where the individual was skipping at the room at an average of 3 skips/s. We collected 50 samples for each activity from a single individual at different timings.

The performance of our system with diverse intensities of ADL is shown in Figure 2(b). Since mild activity is sometimes masked by the multipath effect, therefore, it had a slightly lower performance (*i.e.*, around 96% of detection rate and 3% of false alarm rate). In contrast, the moderate and vigorous ADL attained above 99% of detection rate and less than 3% of false alarm rate.

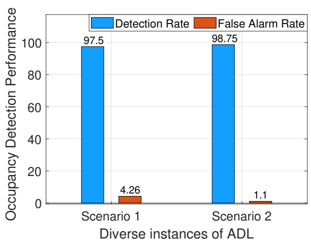
Evaluation of robustness to the instances of ADL. To assess the robustness of our system with variable instances, we considered two scenarios. In the first scenario, we occupied the room for majority of time and then vacate it for shorter time. In the second scenario, we did the opposite, *i.e.*, the room was empty for majority of time, and occupied it for shorter time. To be more precise, we recorded the CSI for 10 minutes in the medium room, where in scenario 1, the room was occupied for 3/4 of the period and unoccupied for 1/4 of the time; vice versa for scenario 2 (*i.e.*, for 8 min unoccupied



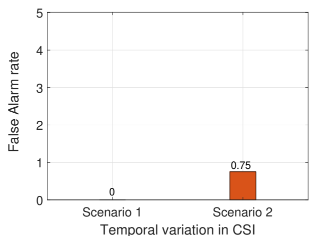
(a) Performance evaluation in enviromental dynamics.



(b) Performance evaluation with the variation of ADL intensities.



(c) Performance evaluation with the variation of ADL instances.



(d) Performance evaluation with temporal variation in CSI.

Figure 2: Robustness evaluation of our occupancy monitoring system with: (a) multipath, (b) variable intensities of physical activity, (c) variable instances of physical activities, and CSI's temporal variation.

and for 2 min occupied). During the occupancy, the moderate ADL was conducted by the occupant.

As evident from Figure 2(c), Scenario 1 has slightly lower performance than Scenario 2. This is because, in scenario 1, there is large cluster for the occupancy state and small cluster for non-occupancy state. As the control limit threshold is derived from the non-occupant cluster, which is smaller in scenario 1 case. In contrast, in scenario 2, there is large amount of non-occupant data, and hence a large cluster for non-occupancy. Thus, the control limit threshold computed in scenario 2 was more precise, thereby scenario 2 slightly outperformed scenario 1.

Evaluation of robustness to the temporal variations in CSI. To evaluate the temporal variation in the CSI measurement, we evaluated our system with two scenarios. First, we took 50 measurements of CSI for a period of for 1 hour at different times in an unoccupied room. The reason for this assessment was to evaluate the robustness of our system to the time varying fluctuations of the CSI in unoccupancy case. In the second scenario, we periodically alternate the occupancy and non-occupancy of the room every 10 minutes for a period of 1 hour. To be more precise, the room was empty for initial 10 minutes, then occupied by the occupant for the next 10 minutes (*i.e.*, 10 $\sim$ 20 minutes), and so forth. In this setting the medium room was considered and moderate ADL was conducted during the occupancy case.

We deliver the temporal variation in CSI in Figure 2(d). We showed only the false alarm rate to focus more on the occurance of temporal variation in static scene (or non-occupancy state). Our system can effective leverage its self-adoptive feature to handle the temporal variation in CSI. In scenario 1, there is no false alarm rate (*i.e.*, false detection of occupant in empty space). Even with alternating the occupancy, the system performs robustly.

4 CONCLUSION

In this work, we devised a self-adoptive system based on off-the-shelf WiFi for occupancy monitoring in smart building to efficiently overcome the energy wastage issue. Our system, does not requires data labeling and pre-training and works robustly in real-time. Moreover, Our system is interoperable to the complex variations in the indoor environments, such as variation in the space size, ADLs having different intensities and instances, and temporal variation in the CSI. Our system attained an average detection rate of 98.9% in diverse indoor environments, 98.5% with diverse intensities of ADL, 98.1% with diverse instances of ADL. Furthermore, our system is also resilient to the temporal variation in the CSI, and achieved less than 1% of false alarm rate.

ACKNOWLEDGEMENT

This research was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education under the grant numbers NRF-2020R1A4A1018774 and NRF-2021M3A9E4080780.

REFERENCES

[1] BBC.com. 2020. *Everyday motion.* Retrieved April 27, 2021 from https://tinyurl.com/2humk9u2

[2] Simona D'Oca, Tianzhen Hong, and Jared Langevin. 2018. The human dimensions of energy use in buildings: A review. *Renewable and Sustainable Energy Reviews* 81 (2018), 731–742.

[3] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *Kdd*, Vol. 96. 226–231.

[4] Michael Hahsler, Matthew Piekenbrock, and Derek Doran. 2019. dbscan: Fast density-based clustering with R. *Journal of Statistical Software* 91, 1 (2019), 1–30.

[5] Robert J Meyers, Eric D Williams, and H Scott Matthews. 2010. Scoping the potential of monitoring and control technologies to reduce energy use in homes. *Energy and buildings* 42, 5 (2010), 563–569.

[6] Sameera Palipana, David Rojas, Piyush Agrawal, and Dirk Pesch. 2018. FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–25.

[7] Vu Minh Quan, Gourab Sen Gupta, and Subhas Mukhopadhyay. 2011. Review of sensors for greenhouse climate monitoring. In *2011 IEEE Sensors Applications Symposium*. IEEE, 112–118.

[8] Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina, and Lucas Rodés-Guirao. 2013. World population growth. *Our world in data* (2013).

[9] Muhammad Salman, Nguyen Dao, Uichin Lee, and Youngtae Noh. 2022. CSI: DeSpy: Enabling Effortless Spy Camera Detection via Passive Sensing of User Activities and Bitrate Variations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–27.

[10] Ville Satopaa, Jeannie Albrecht, David Irwin, and Barath Raghavan. 2011. Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops*. IEEE, 166–171.

[11] Matthias Schulz, Daniel Wegemer, and Matthias Hollick. 2017. Nexmon: The C-based Firmware Patching Framework. https://nexmon.org

[12] Aryan Sharma, Junye Li, Deepak Mishra, Gustavo Batista, and Aruna Seneviratne. 2021. Passive WiFi CSI sensing based machine learning framework for COVID-Safe occupancy monitoring. In *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 1–6.

[13] Elahe Soltanaghaei, Avinash Kalyanaraman, and Kamin Whitehouse. 2017. Peripheral wifi vision: Exploiting multipath reflections for more sensitive human sensing. In *Proceedings of the 4th International on Workshop on Physical Analytics*. 13–18.

[14] Nancy W. Stauffer. 2013. *Reducing wasted energy in commercial buildings.* Retrieved Oct 17, 2022 from https://tinyurl.com/2d2mkccx

[15] NR Tague. 2004. The Quality Toolbox, ASQ (American Society for Quality).

[16] Jianfei Yang, Han Zou, Hao Jiang, and Lihua Xie. 2018. Device-free occupant activity sensing using WiFi-enabled IoT devices for smart homes. *IEEE Internet of Things Journal* 5, 5 (2018), 3991–4002.

[17] Han Zou, Yuxun Zhou, Jianfei Yang, Weixi Gu, Lihua Xie, and Costas Spanos. 2017. Freedetector: Device-free occupancy detection with commodity wifi. In *2017 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*. IEEE, 1–5.

[18] Han Zou, Yuxun Zhou, Jianfei Yang, Weixi Gu, Lihua Xie, and Costas Spanos. 2017. Multiple kernel representation learning for WiFi-based human activity recognition. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 268–274.