

WiFi-Enabled Smart Human Dynamics Monitoring

Xiaonan Guo
Stevens Institute of Technology
NJ, USA 07030
xguo6@stevens.edu

Hongbo Liu
Indiana University-Purdue
University Indianapolis
IN, USA 46202
hl45@iupui.edu

Bo Liu
Stevens Institute of Technology
NJ, USA 07030
bliu11@stevens.edu

Yingying Chen
Stevens Institute of Technology
NJ, USA 07030
yingying.chen@stevens.edu

Cong Shi
Stevens Institute of Technology
NJ, USA 07030
cshi5@stevens.edu

Mooi Choo Chuah
Lehigh University
PA, USA 18015
chuah@cse.lehigh.edu

ABSTRACT

The rapid pace of urbanization and socioeconomic development encourage people to spend more time together and therefore monitoring of human dynamics is of great importance, especially for facilities of elder care and involving multiple activities. Traditional approaches are limited due to their high deployment costs and privacy concerns (e.g., camera-based surveillance or sensor-attachment-based solutions). In this work, we propose to provide a fine-grained comprehensive view of human dynamics using existing WiFi infrastructures often available in many indoor venues. Our approach is low-cost and device-free, which does not require any active human participation. Our system aims to provide smart human dynamics monitoring through participant number estimation, human density estimation and walking speed and direction derivation. A semi-supervised learning approach leveraging the non-linear regression model is developed to significantly reduce training efforts and accommodate different monitoring environments. We further derive participant number and density estimation based on the statistical distribution of Channel State Information (CSI) measurements. In addition, people's walking speed and direction are estimated by using a frequency-based mechanism. Extensive experiments over 12 months demonstrate that our system can perform fine-grained effective human dynamic monitoring with over 90% accuracy in estimating participants number, density, and walking speed and direction at various indoor environments.

ACM Reference format:

Xiaonan Guo, Bo Liu, Cong Shi, Hongbo Liu, Yingying Chen, and Mooi Choo Chuah. 2017. WiFi-Enabled Smart Human Dynamics Monitoring. In *Proceedings of SenSys '17, Delft, Netherlands, November 6–8, 2017*, 13 pages. DOI: 10.1145/3131672.3131692

1 INTRODUCTION

With the rapid pace of urbanization and socioeconomic development, people spend more time gathering in indoor venues for work,

social and physical activities, etc. When multiple people gather together, they exhibit distinct characteristics in terms of crowd size, density distribution and moving patterns that reflect different social behaviors. The awareness of human dynamics is critically important in designing management services [8]. For example, the caring services for seniors, especially for those who have been identified as suffering from social isolation [4] in assisted living places, can be proactively treated once they are identified. In addition, management services in community activity rooms can be planned more effectively if the popularity of different activities within these rooms is understood.

In particular, in assisted living places, due to the gradual loss of physical and cognitive functions, seniors experience many health concerns and they can also easily develop depression [17] [7] caused by social isolation. Thus, it is important to track the activities of seniors in an assisted living place so that their physical and mental health could be assessed in time. For example, tracking the number of elderly people helps to quickly identify whether any elderly people is sitting on the couch without interacting with anybody suffering social isolation in an assisted living place. In other instances, identifying a senior person who is a sleep walker via detecting him walking in a common living room late night or spotting the physically active seniors moving from one activity to another during the day helps to obtain the first-hand information to assess mental state health and social activeness. Additionally, an administrator of a multi-function room (e.g., campus, corporation) often desires to have a comprehensive view of the activities engaged by participants in that room (e.g., watching TV and playing pool table in campus activity rooms; attendance participation of different conference sessions). Such information allows us to infer the most popular activity in a big multi-function room and also aids evacuation planning under emergency situations [9]. To fulfill all these requirements, a human dynamic monitoring system with the following features needs to be designed: (i) estimating the number of participants, (ii) inferring human density distribution, (iii) estimating human walking speed and direction. Toward this end, we propose to reuse existing WiFi infrastructures to perform smart human dynamics monitoring because WiFi technology is prevalent in most indoor venues nowadays and does not require any active human participation.

Recent years have witnessed the emergence of technologies utilizing fine-grained Channel State Information (CSI) in WiFi signals

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SenSys '17, Delft, Netherlands

© 2017 ACM. 978-1-4503-5459-2/17/11...\$15.00

DOI: 10.1145/3131672.3131692

This is the author's manuscript of the article published in final edited form as:

Guo, X., Liu, B., Shi, C., Liu, H., Chen, Y., & Chuah, M. C. (2017). WiFi-Enabled Smart Human Dynamics Monitoring. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems - SenSys '17* (pp. 1–13). Delft, Netherlands: ACM Press. <https://doi.org/10.1145/3131672.3131692>

for people identifications [18] [26]. WifiU [18] presents a CSI-based people identification framework by capturing the unique gait patterns of different people with commercial WiFi devices. Similarly, WiWho [26] studies the gait patterns to perform people identification utilizing various motion impact on CSI amplitudes. Furthermore, WiDir [23] estimates people walking direction leveraging the Fresnel Zones of WiFi signal propagation. The above methods demonstrate the success of human monitoring problem by using CSI. There are also increasing interests of crowd estimation. For example, Electronic Frog Eye [24] performs coarse-grained crowd counting derived from the Verhulst model [12]. However, these studies only focus on one specific aspect of human monitoring. To facilitate the different needs at elderly care facilities and environments involving multiple activities, our work takes a more comprehensive perspective to provide fine-grained monitoring on human dynamics. Specifically, we focus on the density estimation, walking speed and direction derivation for a smaller number of people to facilitate the various needs of elder care and the environments involving multiple activities. For example, our system can first provide a macro view by measuring the number of people and people density. If there is only a single person in a room during day time, we may want to explore if that person is suffering from social isolation and also potentially further monitor his walking speed and direction.

Specifically, we propose an integrated device-free human dynamic monitoring system, which relies on low-cost WiFi devices to infer comprehensive activity-related information leveraging Channel State Information (CSI). In particular, we demonstrate that our proposed system is capable of providing a comprehensive view of human dynamic information in the following three aspects: **(1) Participant estimation.** We develop a semi-supervised learning approach using the non-linear regression method to accurately estimate the number of participant leveraging CSI readings in indoor venues. This method allows our system to be deployed in different locations with significantly reduced training efforts. This information will well serve a variety of public scenarios requiring human dynamic monitoring, such as activity rooms in campuses, corporation meeting rooms, and senior care centers, etc., **(2) Human density estimation.** Our system also proposes a novel approach for the human density estimation based on the analysis of the CSI variance histogram across different WiFi sub-carriers. A broad range of applications would benefit from people density information, such as inferring the most popular activity in a multi-function activity room or social isolation of any senior in an assisted living place. **(3) Walking speed and direction derivation.** Our system develops a frequency-based classification method based on Total Harmonic Distortion (THD) to infer both walking speed and direction. Such detailed information on walking pattern is extremely useful for scenarios where seniors need to be watched closely for their activities to ensure their safety and to provide in-time assistance.

The contributions of our work are summarized as follows:

- We leverage existing WiFi infrastructure to monitor human dynamics without active user participation to facilitate applications of elder care and smart space management. We design the system to perform comprehensive people dynamics monitoring by estimating participant

number and density and inferring walking speed and direction using low-cost commodity WiFi.

- We propose a semi-supervised learning approach to estimate the number of participants. The advantage is that our system only needs to perform limited training, and the constructed model can be generalized and applied to different environments, significantly reducing the labor efforts and make the system scalable.
- We show that our system has the fine-grained capability to estimate participants' density within a monitored area as well as derive people walking speeds and directions to provide first-hand information of seniors' social activeness and smart space requirements.
- Our extensive experiments with 10 people in different types of environments during a period of 12 months demonstrate that our system is highly feasible to estimate the participants' number and density and derive fine-grained walking speed and direction information¹.

The rest of this paper is organized as follows. In Section 2, we review the state-of-the-art solutions and briefly present how the proposed approaches advance the existing ones. We then conduct several experiments to show the insights of our idea in Section 3. We next show our system design overview in Section 4. In Section 5, we present our algorithm design in details. Next, in Section 5.3, we conduct extensive experiments to evaluate our system and show our system is robust under different environments. At last, we conclude our work in Section 7.

2 RELATED WORK

Recent years have witnessed great efforts on exploring a variety of technologies on human monitoring. Traditional approaches are usually categorized into two types, device-based and device-free methods. *Device-based solutions* require assistance of the dedicated devices and sensors (e.g., Wireless devices, dedicated sensors) [1, 3, 10, 16, 20, 21, 27] to monitor human beings. Specifically, Jens *et al.* [21] estimate crowd density by scanning nearby Bluetooth devices with smartphones. Kannan *et al.* [10] perform crowd counting based on audio signal generated and received by the speakers and microphones that on mobile phones. Zihajehzadeh *et al.* [27] derive people walking speed with lower body-mounted Inertial Measurement Unites (IMU). SurroundSense [1] achieves accurate logical localization by sensing the ambient sound, light and color with various built-in smartphone sensors. However, the application scenarios of the above solutions are limited by requiring the users to carry the appropriate equipments. Therefore, *Device-free approaches* are developed to remove such requirement, and utilize either received signal strength from wireless links [14, 25] or

¹The experiments involving human participants have been approved by the authors' institutional ethics review board and necessary consent from all the participants are also recorded.

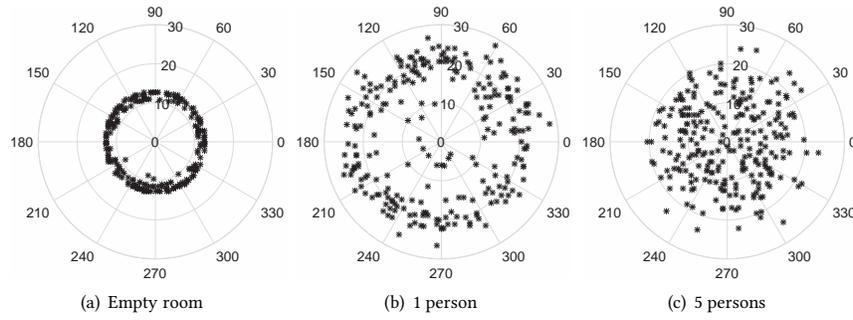


Figure 1: CSI measurements affected by different number of participants: empty room, 1 person and 5 persons.

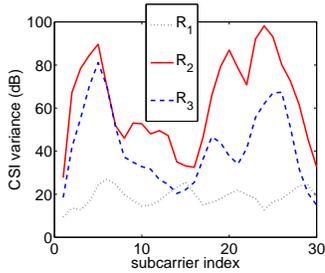


Figure 2: Human density Study: CSI variance when people are in Region 1, 2, or 3 respectively.

image information [2, 13] to facilitate human monitoring. Nakatsuke *et al.* [14] perform passive crowd density estimation by analyzing the Received Signal Strength Indicator (RSSI) and Link Quality Indicator (LQI) obtained from the wireless sensors in indoor environments. Xu *et al.* [25] focus on multi-subject counting and indoor localization by utilizing coarse-grained RSSI along with a successive cancellation algorithm. However, these approaches require extensive deployment of sensor nodes which result in extremely high cost and training efforts. With the assistance of camera-based techniques, Cho *et al.* [2] perform crowd density estimation by studying appropriate features from a sequence of images. Ma *et al.* [13] estimate crowd density through foreground segmentation of video images and then relate the foreground pixel count to the number of people. But the high deployment costs and privacy concerns of these methods hinder them from widespread use.

Unlike traditional RSS-based techniques, which rely on a single feature per transmitted packet, we use Channel State Information (CSI), which provides more fine-grained information (i.e., amplitude and phase) with respect to each OFDM subcarrier per wireless channel. It is readily available in many commercial network interface cards such as Intel 5300 [5] and Atheros 9390 [11]. Nowadays, CSI has been widely applied to various application scenarios especially for *indoor localization*, *activity recognition* as well as *people identification* [6, 18, 19, 22, 26]. To assist indoor localization, Wu *et al.* [22] use CSI to detect the Line-Of-Sight path in an indoor environment. They studied the correlation between CSI streams and estimate Angle of Arrival (AoA) based on phase differences between different antennas. E-eyes [19] performs in-place activity recognition by utilizing CSI measurements. This work uses Dynamic Time Warping (DTW) to classify walking activities and use Earth

Mover Distance (EMD) for similarity comparison. WiFall [6] was proposed to detect the falling activity for individual person based on CSI measurements. They extract features from CSI streams and develop a one-class Support Vector Machine (SVM) to differentiate falling activity from the others. In the direction of people identification [18] [26], WiFiU [18] identifies human by analyzing frequency spectrum of WiFi traffic to derive the walking behavior and gait. Similarly, WiWho [26] focuses on people identification by utilizing CSI amplitude to identify a person’s steps and walking gait. A peak-and-valley method is developed to identify step cycles for effective feature extraction. However, the aforementioned works only focus on one specific aspect (i.e., localization, activity recognition or human identification) on human dynamics, so a systematic study on monitoring human dynamics is desired.

Along this trend, there are works which focus on crowd counting [24] and walking direction estimation [23] by leveraging CSI-based techniques. WiDir [23] proposes to estimate the human walking direction based on a Fresnel Zone based model which depicts the relationship between walking direction and CSI phase shift. However, a dedicated device setup (i.e., at least three laptops to form two Fresnel zones) is required to ensure that the system can operate properly. Electronic Frog Eye [24] performs coarse-grained crowd counting by using the Verhulst model [12], which maps a specific CSI variance level to the corresponding number of walking people. This study only provides coarse-grained crowd counting information and can not be deployed to infer other critical human dynamic information (i.e., people density and walking speed). Our work, from a different angle, aims to provide fine-grained monitoring of human dynamics by studying a small number of people with density estimation and walking speed and direction derivation, which are suitable for mental state health and social activity study in elder care and efficiency activity management in multi-purpose facilities.

3 PRELIMINARY

3.1 CSI Fundamental

In order to facilitate the human dynamics monitoring by leveraging CSI measurements in WiFi networks, we first need to understand how human dynamics affect the wireless channel. In a multi-path wireless environment, the received wireless signal can be expressed as: $y = Hx + n = e^{j\omega_0 t} \sum_{i=1}^N A_i e^{-j\omega_0 \tau_i} + n$, where y and $x = e^{j\omega_0 t}$ are the receiving and transmitting signals in vector form, respectively, and $H = \sum_{i=1}^N A_i e^{-j\omega_0 \tau_i}$ and n are

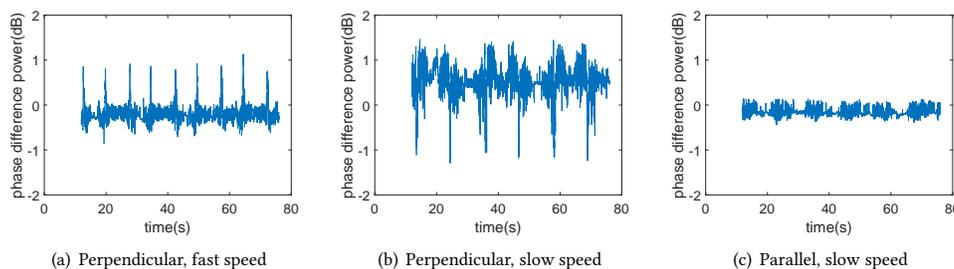


Figure 3: Relative phase variation when walking in perpendicular or parallel to the wireless link with different walking speeds.

the channel response and the ambient noise. Specifically, A_i is the complex amplitude attenuation for the i^{th} , $i \in (1, N)$, path as $a_i e^{j\phi_i}$, $e^{-j\omega_0 \tau_i}$ is the phase shift on the i^{th} path with a propagation delay τ_i . Note that, $\tau_i = \frac{l_i}{c}$, where l_i is the length of the i^{th} path and c is the speed of light.

The estimation of H represents the channel state information (CSI), which is readily available from many commercial wireless devices. \hat{H} can be expressed as follows: $\hat{H} = \sum_{i=1}^N A_i e^{-j\omega_0 \frac{l_i}{c}}$. This equation implies that \hat{H} is dominated by three major factors, *amplitude attenuation*, *phase shift* and the *propagation delay*. The presence of people and their movement would result in different CSI variations in the area of interest. In particular, the signal strength is attenuated when the line-of-sight (LOS) path is blocked by human body, and human body changes the number of propagation paths as well as introduces more reflections that alter the phase at the receiver side. Based on the analysis above, we find the amplitude and phase variations on Channel State Information (CSI) could reveal the fine-grained information on people movements. Integrating the CSI measurements across multiple subcarriers, it is possible to infer the human dynamics via statistical analysis.

3.2 Feasibility Study

3.2.1 Experimental Methodology. In order to utilize CSI measurements to depict the human dynamics (i.e., the number of participants, human density, walking speed and direction), it is necessary to understand the relationship between different human dynamic patterns and the corresponding CSI measurements. Towards this end, we conduct a set of preliminary studies in a typical classroom of the size $12m \times 5m$, where several chairs and tables are placed. Two laptops, acting as a pair of transmitter and receiver, are placed at the two ends of the classroom and hence they are 5m apart from each other. The packet sending rate is fixed at $100pkt/sec$, and 5 volunteers participated in this experimental study.

3.2.2 Impact of the Number of People on CSI. In this part, we show that the CSI measurements will certainly be affected by the presence of people in a room due to their impacts on the wireless signal propagation paths (i.e. reflection and refraction). Intuitively, as the number of people increases, more signal propagation paths will be affected, resulting in different CSI readings. Figure 1 (a)-(c) plots the raw CSI values against polar coordinates, which includes both the amplitude and phase information lasting for 30 seconds

under the scenarios with no person, 1 walking person and 5 walking people, respectively. It is obvious that the CSI amplitude readings are quite stable in an empty room. On the contrary, when people are moving inside the room, both the CSI amplitude and phase are affected differently with varying number of people. This result motivates us to explore the relationship between the CSI reading and the number of participants. After we obtain such information, we can further quickly identify the situation where one elderly person wanders away.

3.2.3 Impact of the Human Density on CSI. Estimating human density is also important since it allows us to infer some important context information, e.g. the most popular activity in a big multi-functional room or the most popular display in a big museum display room. Here, we show that different people densities inside a room should also have different impacts on the CSI readings. The rationale lies in the fact that when more people gather at some specific regions, the wireless signal will experience heavier distortions caused by their movements in those regions. This could help us on differentiating the people density via CSI readings. We use one pair of laptop and ask 3 volunteers to move at three different regions. The three different regions are defined as follow: the area close to the transmitter; the area close to the receiver; the area in the middle of the transmitter/receiver pair. We collect a minute of CSI readings for each case and analyze the collected data. The results are plotted in Figure 2. The results show that the distribution of CSI variance across different subcarriers for different people densities in the room. It inspires us to utilize the CSI distribution to perform people density estimation.

3.2.4 Impact of the People Walking Speed and Direction on CSI. Walking speed and direction are critical information which allows us to infer human dynamics, e.g., inferring if a senior walks out of his bedroom to a common living room at night. In order to find out how the CSI readings are affected by different walking speeds and directions, we ask the a volunteer to walk back and forth between two ends of a room with different speeds and directions while we collect CSI readings on both wireless links. As shown in Figure 3, the CSI readings have significant differences as the walking speed and direction change, and the peaks correspond to the moment when walks across the LOS of wireless link. Specifically, slow walking speed causes larger variation than fast walking speed on the power of phase difference, while walking along the LOS of wireless links has less impact on the CSI readings than walking across

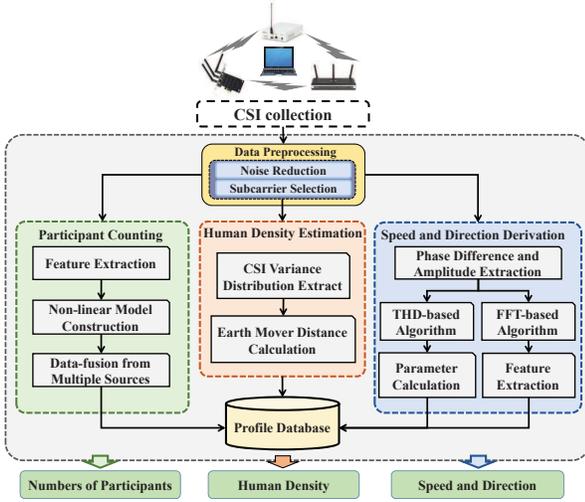


Figure 4: System overview of human dynamic monitoring.

the LOS. Inspired by the above observations, we will study the relationship between the power distributions of phase difference in frequency domain and walking speed & direction.

4 SYSTEM DESIGN

The proposed system aims to develop a systematic approach to enable comprehensive human dynamics monitoring to assist smart management in public indoor environments. To achieve such a goal, we propose to utilize the readily available CSI information in commodity wireless devices to sense how the human dynamics affect the wireless channel.

The architecture of our proposed system is shown in Figure 4. In general, our system will first monitor the macro perspective of an area of interests through estimating the number of participants and human density. Then, if only one person is present in the monitoring area, we will have a closer look at this person by measuring the micro perspective via walking speed and direction derivation. First, our system takes time-series of raw CSI readings as the input and each CSI measurement contains $30 N_t \times N_r$ matrices, where N_t and N_r are the number of antennas on transmitter and receiver, respectively. We first feed the raw CSI readings into the *Data Preprocessing* module to filter out high frequency noise that are caused by interference or ambient noise. Next, the *Subcarrier Selection* module will remove the subcarriers that are not sensitive to capture human dynamic characteristics. Then, the preprocessed CSI readings will be fed into three functional modules, *Participant Counting*, *Human Density Estimation*, *Speed and Direction Derivation*. In particular, the *Participant Counting* module performs quantitative analysis of the CSI readings to derive the number of people in a specific area; the *Human Density Estimation* module estimates the human density information based on CSI variances over subcarriers; and the *Speed and Direction Derivation* module classifies the walking speed into different levels and estimate the walking direction based on the frequency-based analysis.

In the *Participant Counting* module, we first investigate and determine the effective features that has monotonic relationship with

the number of participant capturing the human dynamic characteristics from the CSI readings, and then describe the monotonic relationship via a non-linear model. During the modeling phase, the system integrates the CSI features from multiple dimensions, including different links (i.e., four laptops form two orthogonal links), frequency bands (i.e., 2.4GHz and 5GHz) and channels (i.e., our system scans the adjacent orthogonal channels and concatenate them together). Given the model built upon one specific room, it can be generalized and applied to other indoor environments without repeating the whole training process. Specifically, we only need to extract the features from another empty room via opportunistic approaches (e.g., collecting CSI measurements at mid-night or early morning when no one is there), and then the non-linear model will update the critical parameters based on these new features to perform participant counting in a new environment.

Human Density Estimation module aims to provide human density information inside the monitoring area, which would assist various application scenarios as described in Section 1. Specifically, it consists of two components: *Distribution Analysis over Multiple Subcarriers* and *Earth Mover Distance Calculation*. Our system first constructs the human density profiles based on the CSI amplitude distribution over multiple subcarriers under different densities in the monitoring area. Next, given a new time series of CSI, we calculate its CSI amplitude distribution and then determine the current density by examining its Earth Mover Distance extracted from the predefined profiles.

Three submodules are involved in deriving the speed and direction information, namely *Phase Difference Information Extraction*, *Total Harmonic Distortion (THD)-based algorithm* and *Estimation parameter calculation*. Given the relative phase information extracted from the raw CSI measurements, we develop the *THD-based algorithm* and *FFT-based algorithm* for walking speed and direction estimation, respectively. Specifically, THD-based algorithm estimates the walking speed based on the fundamental signal extracted from the relative phases through THD analysis, while the FFT-based algorithm determines the walking direction by classifying the frequency-domain features of the relative phases into one of the predefined profiles with respect to different walking directions.

5 ALGORITHM DESIGN

In this section, we first present our participant number estimation and human density estimation scheme, and then propose the new approaches for deriving people walking speed and direction.

5.1 Semi-Supervised Participant Number Estimation

Estimating the number of participants is of great importance especially in providing smart monitoring in senior care or multi-functional room scenarios. To achieve this goal, we propose a semi-supervised learning algorithm to derive the monotonic relationship between the number of people and relevant statistical features extracted from CSI measurements.

5.1.1 Feature Extraction. Unlike many existing works, which rely on the CSI measurements collected from a single antenna, we take the advantage of the spatial diversity on the MIMO-enabled

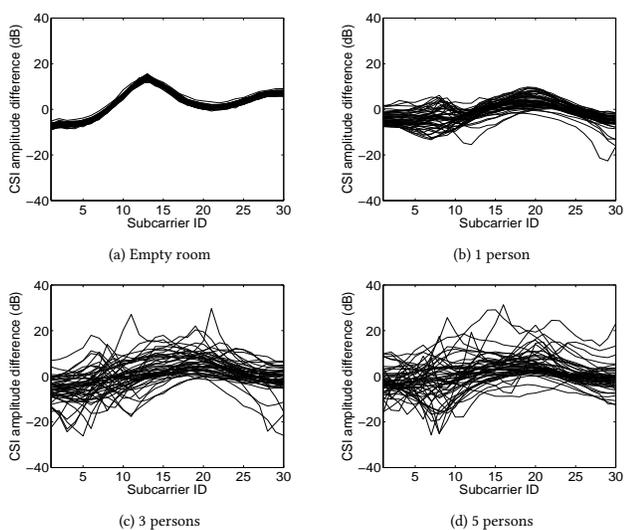


Figure 5: CSI amplitude difference between two antennas under different number of participants.

WiFi devices to extract effective features from the CSI difference between two antennas. As the wireless channel is easily affected by various factors (i.e., hardware instability, environmental changes, etc.), resulting in unpredictable offset of CSI readings on a single antenna, given the fixed distance between two antennas, the CSI differences will only retain the human dynamics related information.

Figure 5 shows the impact of the number of participants on CSI amplitude difference with no person, 1 person, 3 persons and 5 persons, respectively. We examine several features extracted from the amplitude and phase information of CSI measurements, and finally determine the following 4 features that follow a monotonic relationship with the number of participants: *CSI amplitude variance*, *CSI amplitude range*, *CSI amplitude mode*, *Entropy of HIP* (Hermitian Inner Product) (i.e., *phase difference*). Figure 6 shows the monotonic varying trend of the four different features when there are 1 to 6 people in the monitoring area, respectively. In the proposed system, we calculate these features for each of the selected subcarriers and then average the feature values to reduce the computational complexity.

5.1.2 Feature-based Non-linear Regression. To derive the monotonic relationship between the number of participants and relevant features, we formulate the relationship as a non-linear regression model. Specifically, we adopt the logistic bounded exponential function whose growth rate is proportional to the size of the function's value and the growth is bounded by some fixed capacities. This model has been widely used for modeling resource limited exponential growth such as island populations, bacteria in a petri dish, and growth of seedling. Similarly, the space of the monitored area is limited and thereby physically restricts human's walking range. In other words, the impact of the number of people is also bounded and the values of the relevant CSI-based features have finite increasing range as the number of people increases. The logistic bounded exponential model is expressed as:

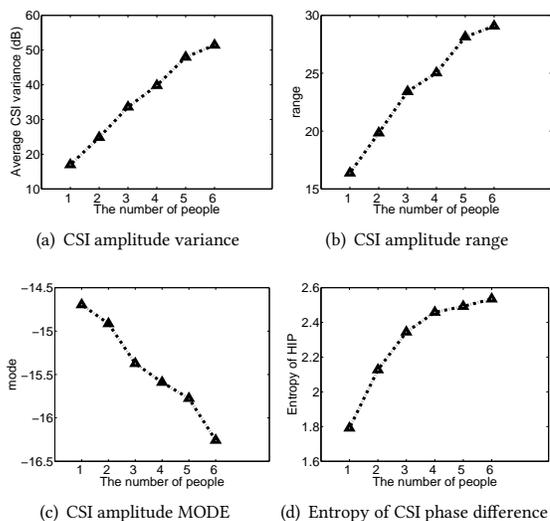


Figure 6: Monotonic relationships between different features and the number of participants for human density estimation.

$$f(x) = \frac{a}{b + ce^{-dx}}, \quad (1)$$

where x is the number of people and a, b, c, d are the four parameters that control the growth rate and range of the function $f(x)$. We can express the number of people as a function of these four parameters via the inverse function $f^{-1}(x)$. Note that it is only necessary to derive useful values for these 4 parameters once at a certain location. During this training process, the relationships between parameters c and d can be identified and hence one only needs to estimate the new a and b values for subsequent deployment of our system at a new location. This is because the shape and the offset of the model are determined by the parameters a and b while parameters c and d depict the shape curvature which is relatively stable across different environments.

The values of the parameters in Equation 1 are derived during the construction of our non-linear model. Specifically, we first collect CSI measurements under five scenarios, including an empty room with no people inside the monitoring area, and 1 to 4 people walking inside the area. Then, we extract the aforementioned features as discussed in Section 5.1.1. For each feature, we have several instances associated with different number of participants and we further fit the feature values into Equation 1 to estimate the parameters by solving a corresponding objective function.

After obtaining all single-feature non-linear models, the final estimate on the number of participants is obtained by minimizing the total estimation errors produced by these N features as follow:

$$\begin{aligned} & \underset{\hat{y}, w_i}{\text{minimize}} && \sum_{i=1}^N w_i (y_{f_i} - \hat{y})^2 \\ & \text{subject to} && \min(y_{f_1}, \dots, y_{f_i}) \leq \hat{y} \leq \max(y_{f_1}, \dots, y_{f_i}), \\ & && \sum_{i=1}^N w_i = 1, \end{aligned} \quad (2)$$

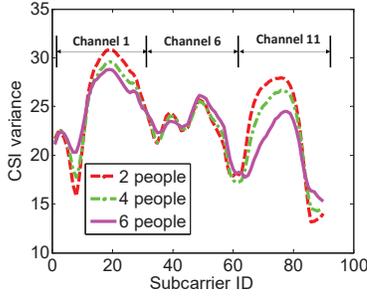


Figure 7: Concatenation of multiple channels with different number of participants.

where y_{f_i} is the output from the i^{th} feature and \hat{y} is the fused estimate, w_i is the weight associated with its error and can be determined via experiments.

After the model is constructed, it can be easily extended to other environments via a semi-supervised learning approach. Specifically, the proposed non-linear model in Equation 1 relies on the parameter a and b to determine the trend of CSI features, which are similar at different locations, as the number of people increases. Therefore, we can determine such similar trend in an empty room to estimate these two parameters via opportunistic approaches, i.e., collecting CSI measurements at mid-night or the early morning when there is no people presence. The other two parameters c and d , depicting the curvature of the model, usually has little impact of the trend. Thus, our system can be easily adapted to run at different locations with little training efforts. Further, our system uses the CSI measurement in the empty room to update the model parameters for the new environment, which is equivalent to an equation solving process.

5.1.3 Data Fusion. To ensure the robustness and high accuracy on participant number estimation, we fuse not only multiple features but also multiple links and multiple channels.

Multiple Links. Multiple transceiver links provide redundant dimensions on CSI collection to eliminate the unstable readings and outliers, so that a more robust estimation result will be ensured in our system as follows:

$$F_L = \sum_{i=1}^{N_L} w_{L_i}(F_{L_i}), \quad (3)$$

where, F_L is the fusion results from multiple links, F_{L_i} is the output from the i^{th} link with the associated empirical weight w_{L_i} and N_L is the total number of links. Note that $\sum_{i=1}^{N_L} w_{L_i} = 1$.

Multiple Channels. For each transmission link, we choose three adjacent orthogonal channels for feature extraction. Our prototype system can switch automatically among these 3 channels and hence the CSI readings of a total of 90 subcarriers will be collected simultaneously instead of 30 subcarriers from one single channel as in existing works. Figure 7 illustrates an example explained as follow: we collect CSI from three different orthogonal channels (i.e., channel 1, channel 6 and channel 11) at 2.4GHz with 2 people, 4 people, and 6 people walking inside the room, respectively. We observe a continuous CSI amplitude variance across all 90 subcarriers and therefore we can perform subcarrier selection

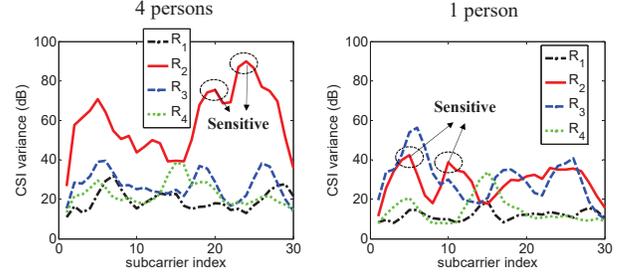


Figure 8: Similar patterns of CSI variance under different number of participants for human density estimation.

with high flexibility. The channel concatenation does not only provide high dimensional CSI measurement from more subcarriers, but also has the capability to resist possible interference over some subcarriers. The final output is obtained as:

$$F_C = \sum_{i=1}^{N_C} w_{C_i}(F_{C_i}), \quad (4)$$

where F_C is the fusion results from multiple channels, F_{C_i} is the output from the i^{th} channel with the associated empirical weight w_{C_i} and N_C is the total number of channels. Note that $\sum_{i=1}^{N_C} w_{C_i} = 1$. Combine Equation 3 and Equation 4, we have:

$$F = \sum_{j=1}^{N_L} \sum_{i=1}^{N_C} w_{L_j}(w_{C_i}(F_{C_i})), \quad (5)$$

where F is the fusion results based on both multiple links and multiple channels.

In particular, our prototype sets the weight of each wireless link as $W_L = 0.5$ since there are two wireless links deployed in our system. In addition, we set the weight of each channel as $W_C = 0.25$, since our prototype employs 4 different channels for participant estimation, where three of them work at 5GHz, and the remaining one is at 2.4GHz. Note that the reason why more 5GHz channels are used is that 5GHz frequency band experiences less interference than 2.4GHz.

5.2 Human Density Estimation

Beside participant counting, our proposed system also estimates the people density distribution within the area under surveillance. People density estimation is of great importance since it allows us to infer the denser areas as well as other general crowd information which facilitate space management and seniors' health service planning. For example, we could infer the popularity of the activities in a multi-function room based on the density information and accordingly expand on the space housing the most popular activity. To understand how different people density distributions affect CSI measurements, we first define the densest region as follow: the region that contains more people than other regions will be denoted as the densest region.

Take two perpendicular transceiver links as an example, it divides the monitoring area into four regions. We can further define the four regions as: Close to Rx_1 and Rx_2 (R_1); Close to Rx_1 and Tx_2 (R_2); Close to Tx_1 and Tx_2 (R_3); Close to Tx_1 and Rx_2 (R_4); Then,

we construct the profile database. In particular, for each of the sub-region, we build a corresponding profile with 4 people. To infer the density region at the monitoring phase, our system collects CSI and calculates the variance distribution across all subcarriers. The system will perform similarity matching by applying the Earth Mover Distance (EMD) [15], which is a well-known approach for evaluating the similarity between two probability distributions, to the CSI variance distribution obtained in a sliding time window. In particular, it calculates the minimal cost to transform one distribution into the other. Our system seeks to derive the EMD by comparing the CSI variance distribution of the testing CSI measurements with the profiles of known human density region. CSI measurements being tested are identified to one of the profile with the minimal cost (i.e., minimal EMD distance).

We further seek solutions to reduce the labor-extensive profiling process. We find that the human density information is preserved in the distribution of the CSI variance across all subcarriers. That means the profiles built from a fixed number of people would be enough. The intuition behind this lies in an experimental observation as shown in Figure 8. We plot the difference of CSI variance from two antennas against 30 subcarriers when volunteers are at four different regions with 1 person and 4 persons respectively. We find the presence of people at different regions leads to a different distribution of the CSI amplitude variance. We can define the *sensitive subcarrier* whose neighbors have lower variance as marked in the figure. We further find that the sensitive subcarriers are location-dependent, regardless of the number of people and only the value is different. Consequently, the profiles built from 1 person have the same series of sensitive subcarriers as the profiles built from 4 persons. This further implies that the human density information is preserved in the distribution of the CSI variance across all subcarriers. The proposed system can take advantage of this property to reduce the training size by only constructing the profile based on a fixed number of people in an environment.

5.3 Walking Speed Derivation

The walking speed of people is a critical descriptor of human dynamics, and essential for human dynamics monitoring and smart management under possible emergencies. For example, walking too fast or slow may indicate some abnormal activities, especially for the senior people. The changes in walking speed would result in different patterns of CSI readings. In this subsection, we first present our walking speed derivation scheme leveraging the CSI measurements.

In our preliminary study, given the spectrum derived from the time series of CSI relative phases, we find the power level in lower frequency band decreases as the working speed increases. Therefore, a total harmonic distortion (THD) based method is developed to achieve accurate speed estimation from CSI readings. THD is traditionally used to characterize the linearity of audio systems and the power quality of electric power systems. It can also help to identify the linear relationship between the fundamental signal embedded in the relative phase information and the walking speed regardless any people.

Given the relative phase from the extracted CSI samples, we first subtract the mean value from the relative phase across 30 subcarriers. Next, we need to identify the periods that capture the walking

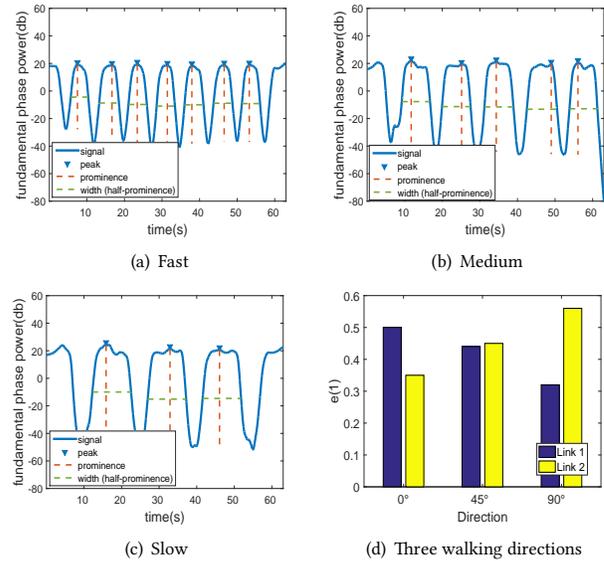


Figure 9: (a), (b) and (c) are THD fundamental signal power from three different walking speeds. (d) is the energy distribution of the first bin from two transceiver links for walking direction estimation.

characteristics of people in CSI samples. Specifically, a hypothesis test is adopted to compare the CSI energy level within a sliding window (i.e., 1sec) with a predefined threshold to determine whether there is anyone walking in the area of interest. If the energy level is constantly below the threshold, it is marked as a static period; otherwise walking period. We apply both Butterworth filter and moving average filter to the CSI relative phase in static period, so that the environmental impact, including the high-frequency noise and outliers, will be effectively removed.

Once the above preprocessing on raw CSI measurements is completed, we then adopt THD analysis to perform walking speed estimation. THD analysis derives the periodogram of input signals using a Kaiser window with large sidelobe attenuation. Our speed estimation scheme searches for the largest nonzero spectral component, called fundamental frequency, across the periodogram, and then computes the central moment of all adjacent bins that decrease monotonically away from the maximum. The walking steps of people always create periodic impacts on the wireless channel, which corresponds to the amplitude of fundamental frequency in THD analysis and implies the walking speed information. The higher order frequencies in THD analysis reflect the arms movement or the environment noise during the walking period, and will be removed for better walking speed estimation. Given a sliding window of 1sec length applied to the CSI readings, our THD-based method extracts the fundamental signal to reconstruct a new time series of CSI readings. As the example shown in Figure 9 (a-c), a volunteer walked several rounds between two ends of a classroom at 3 different speed levels, and stopped at both ends for a few seconds to separate the walking and static periods during each round. It is obvious to find that the prominence and width of fundamental signals for each walking round increase as the walking speed decreases.

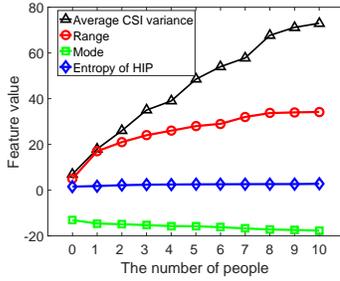


Figure 10: Monotonic relationship between features and the number of people.

According to our empirical study, each peak of the fundamental signal stands for the scenario that the person walks across the LOS between the transmitter and receiver. The signal power decreases as the person walks away, and vice versa. Taking a close look at the fundamental signal variation, we find higher walking speed causes smaller fluctuation on the fundamental signal, resulting in small prominence h around the peak of the fundamental signal due to the shorter dwelling period at each position. Further, the width w at half prominence h implies the time taken to walk the same distance. Intuitively, higher walking speed has smaller w . Given both the prominence h and width w information, we could derive the estimation parameter as $e = h \times w$, which has a monotonic relationship with walking speed.

We tested a few different walking routes in different rooms, our algorithm can estimate the walking speed if it is relatively stable in different locations. Even though different scenarios may have impacts on the CSI readings, our algorithm can still derive the estimation parameter e to estimate the walking speed. It is possible that the algorithm may fail if the walking speed changes dramatically in a short time window (e.g. 1s) which we will explore in our near future work.

5.4 Walking Direction Estimation

Walking direction estimation can be useful in scenarios related to senior care services. In particular, we focus on the walking direction of a single people, especially for the sleep walker, who may walk towards dangerous places (e.g., staircases or balcony) at night.

To derive the walking direction, we first segment CSI measurements using a fixed time window (3 ~ 5 seconds) and define the frequency distribution as: $e(n) = \frac{f_n}{\sum_i f_i}$, where $e(n)$ represents the frequency distribution of the n th output f_n from the FFT operation, and f_n is frequency bin (i.e., sub-band). For example, given the sampling rate at 50Hz, a 5-points FFT transform converts the frequency band of 50Hz into 5 bins with each representing 10Hz bandwidth. To further demonstrate our algorithm, we ask a volunteer to walk in three directions: parallel to link 2 (0°); walking across both links (45°); and parallel to link 1 (90°). In the meanwhile, the CSI over both links are collected to examine their frequency distribution. As shown in in Figure 9 (d), we only examine the dominant frequency bin $e(1)$, which locates at the lower frequency band, for both links on three walking directions, as it is affected more by direction changes than the other frequency bins. We found that, when the person walks parallel to link 1, $e(1)$ on

this wireless link has lower energy level than that on the other wireless link 2 perpendicular to link 1. Further, if the the people walks along the 45 degree direction, which crosses two wireless links, $e(1)$ on both links have similar energy level.

We also utilize CSI variance to enhance direction estimation performance. In general, given a fixed time period (i.e., 4 seconds), fast speed in one direction causes low CSI variance due to the short duration at each location, whereas low speed causes high CSI variation. Inspired by two above interesting observations, we propose to leverage both CSI frequency distribution and variance on two perpendicular wireless links as a combined feature vector, z , to estimate walking direction. Our system trains a support vector machine (SVM) model with Gaussian kernel and classifies the walking direction based on the pre-built walking direction profiles above, which include three walking directions (i.e., 0° , 45° , 90°). The SVM model then classifies the combined features to the closest walking direction. More precise direction estimation can be achieved with fine-grained direction profiles.

Currently, we focus on estimating a single person's walking direction, and we will further investigate the capabilities of our system on dealing with the cases when multiple people walk together and have different walking directions in the near future.

6 PERFORMANCE EVALUATION

6.1 System Prototype

We implement our system on two pairs of laptops equipped with Intel 5300 802.11n WiFi NIC acting as WiFi transceivers. Specifically, the laptops form two transceiver links with their LOS are perpendicular to each other. The laptops for each pair of transceivers, we let the Lenovo T500 laptop as receiver, while Dell E6430 laptop acts as the transmitter. Both transmitters are configured to transmit packets at the packet rate 100 pkt/sec, but work at different WiFi channels to avoid conflicts. The receivers work at Monitor mode, which enables the receivers to capture and decode the 802.11 frames sent by the transmitters. The CSI values are extracted from these frames with the CSI tool developed by Halperin *et al.* [5]. The captured CSI measurements contain the samples from 30 subcarriers, and each subcarrier entry is a complex number with signed 8-bit resolution for both the real and imaginary parts. The prototyped system can work at both 2.4GHz and 5GHz frequency bands involving multiple 20MHz bandwidth channels.

6.2 Experimental Methodology

We conduct extensive experiments by recruiting 10 volunteers to move at 5 different indoor venues over a 12-month time period. Specifically, The experiments are carried out twice per week with each lasting for 2 to 3 hours, and the total amount of CSI data collected during the experiments is about 10GB. Each time before our experiments, we detect the channel occupancy by opportunistically scanning the channel usage at both 2.4GHz and 5GHz. Then, we select the channels that experience less interferences in our system. In addition, given the fixed number of people in the area of interest, we find that irrespective of the combinations of people with different sizes and dresses, the CSI readings are less impacted. Furthermore, we also study the impact of different placements of

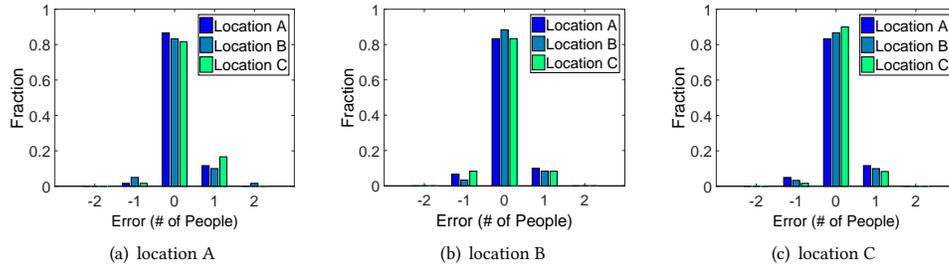


Figure 11: Performance of participant estimation by using CSI in empty room of location A, B, C for training.

APs for collecting the CSI measurements, and still obtain consistent results. Through the above experiments, we demonstrate that a comprehensive view of human dynamics could be obtained leveraging existing WiFi infrastructures.

Participants and Human Density Estimation. To evaluate the performance of the number of participants and human density estimation, we design and conduct the experiments at three different indoor scenarios (i.e., activities room (Location A), meeting room (Location B) and classroom (Location C) with dimensions $16m \times 10m$, $13m \times 7m$ and $12m \times 5m$, respectively). For the three environments, we observe that the people moving in the next rooms will not change our results since the CSI is mainly affected by the number of participants in the current room. In total, we collect and process over 100 hours of CSI measurements and there are 10 volunteers participate in our experiments.

Walking Speed and Direction Derivation. To evaluate the performance of THD-based speed estimation algorithm, we collect 1296 different walking traces of 4 people in two different classroom environments of size $7m \times 6m$ (big classroom) and $12m \times 5m$ (small classroom), respectively. Each walking trace involves 3 different speed levels: fast ($1.33m/sec$), medium ($1m/sec$) and slow ($0.67m/sec$), and each speed level spans a certain period on the walking trace. To evaluate the performance of our classification-based walking direction estimation, we collect 286 experimental samples of 3 people in the big classroom. In particular, each person is asked to walk in 3 different directions (i.e., 0° , 45° , 90°). The experiment processes are recorded by a camcorder serving as ground truth for performance evaluation.

6.3 Evaluation Metrics

To quantify the estimation accuracy for each module in our system, we define the following metrics for performance evaluation:

Participant Estimation Accuracy. To evaluate the performance of our proposed semi-supervised learning approach, we define the *Participant Estimation Accuracy* as the ratio between the correct instances of participant estimation and the total number of instances.

Human Density Accuracy. To evaluate the performance of human density estimation scheme, our system profiles the CSI patterns when the volunteers walk inside different regions with various people densities in the area of interest. Then the human density accuracy is defined as the ratio between the number of correctly profiles matching and the total number of experiments.

Walking Speed and Direction Accuracy. The proposed system categorizes 3 speed levels (i.e., slow, medium, fast) and 3 walking directions (i.e., 0 degree, 45 degree and 90 degree deviated from

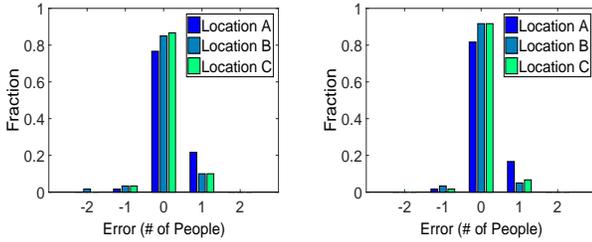
the LOS of one wireless link). We define the speed and direction accuracy as the percentage of the correct classification results.

6.4 The Number of Participants Estimation

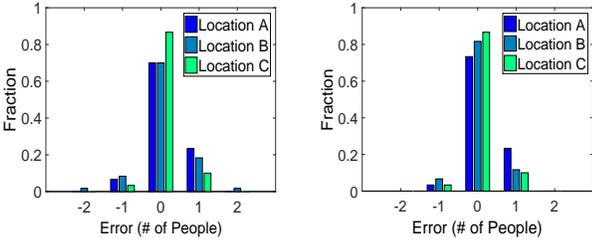
6.4.1 Non-linear Model Construction. To build an accurate participant estimation model, it is critical to correctly identify the relationship between the number of people and the features extracted from CSI readings. More specifically, our system seeks to explore such a monotonic relationship to easily and accurately predict the number of people based on CSI readings. We first demonstrate that the CSI features we selected exhibit a monotonic relationship with the number of people. Figure.10 shows the four selected features versus the number of people from 0 (i.e., empty room) to 10. It is obvious that all the selected features (i.e., average CSI variance, Range, Mode and Entropy of HIP) as introduced in Section 5 monotonically increases or decreases as the number of people increases.

6.4.2 Semi-supervised Learning Approach. Our proposed semi-supervised learning approach could build a generalized model to apply limited training data to different environments without introducing extra training effort. Particularly, given the CSI readings collected from a new empty room, the model will be automatically updated to fit into the new environment for the estimation of the number of participants. Figure 11 shows the estimation accuracy with the model built upon the CSI profiles at the three different rooms. All the rooms have high accuracy on estimating the number of participants (i.e., 90%, 89% and 91%) when building the model based on the CSI profiles at the same room. When the model in room A is built based on the CSI profiles collected at different rooms (i.e., B and C), we observe that the estimation accuracy still maintains as high as 83% and 81%, respectively. Similarly, the model at location B achieves the estimation accuracy 84% and 85% based on the profile from A and C, while the model at location C has 85% and 89%. Overall, all the three locations show the estimation accuracy over 80% even without training, which proves the proposed system can achieve high accuracy on estimating the number of participants with the proposed semi-supervised learning approach.

6.4.3 Data Fusion. Next, we study the impact of data fusion from different dimensions (i.e., multiple frequency bands, multiple channels) of CSI measurement on the performance of the proposed system. **Multiple Frequency Bands.** First, we evaluate the performance of our system when working at different frequency bands. Figure 12(a) and (b) show the participant number estimation results at 2.4GHz and 5GHz, respectively. For three different



(a) 2.4G for all three locations with channel combination (b) 5G for all three locations with channel combination



(c) 2.4G for all three locations without channel combination (d) 5G for all three locations without channel combination

Figure 12: Performance of participant counting at 2.4GHz and 5GHz band with and without channel combination respectively.

locations, the overall accuracy at 5GHz outperforms that at 2.4GHz. In particular, the accuracy at location A, B and C are 78%, 84% and 86% for 2.4GHz, while the accuracy are 82%, 92% and 92% for 5GHz. This is caused by the poor link quality at 2.4GHz frequency band, which involves more collisions among many WiFi transmissions. The results demonstrate that our system can maintain good performance by only operating at different frequency bands without much accuracy degradation.

Multiple Channels. We also examine the impact of different channel combinations on the participant number estimation in our system. Particularly, the system combines the three adjacent orthogonal channels (i.e., channel 1, 6, 11 at 2.4GHz and channel 40, 44, 48 at 5GHz). Thus, we provide redundancy information since we have 90 subcarrier candidates to choose from compared with 30 subcarrier from one channel only. Such channel concatenation not only provides high dimensional CSI readings from more subcarriers, but also has the capability to resist possible interference to the CSI readings over some subcarriers. Figure 12 shows the experimental results with different channel combinations given in Figure 12(a) and (b) and without channel combination given in Figure 12(c) and (d). We observe that without channel combination, the system performance at three different locations degrades to 70%, 70% and 86% at 2.4GHz and 74%, 81% and 87% at 5GHz. The results prove that using more channels improves the performance of our system up to 8%.

6.5 Human Density Estimation

In this subsection, we study the performance on human density estimation in our system under different people distributions.

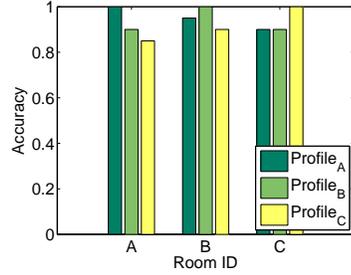


Figure 13: Performance of human density estimation. $Profile_A$, $Profile_B$ and $Profile_C$ indicate the profiles are built from location A, B and C, respectively.

Table 1: EMD distance at location A

Regions	R_1	R_2	R_3	R_4
R_1	3.16	4.76	6.14	6.41
R_2	4.28	1.69	3.75	8.28
R_3	6.03	3.78	2.29	10.74
R_4	6.13	8.66	11.28	1.97

Different Indoor Environments. To evaluate the performance of people density estimation, we conduct experiments at three different indoor environments. We first divide the space of each indoor environment into four regions (i.e., R_1 , R_2 , R_3 and R_4) based on two links, and the volunteers are asked to move at different regions forming different density distributions. Meanwhile, the corresponding CSI measurements (i.e., testing data) will be collected to compare with the predefined CSI profiles (i.e., training data) of different density distributions. We try in total 20 different density distributions and obtain the overall performance shown in Figure 13. We observe that the proposed system achieves 100% accuracy on determining the people density distribution in the room based on the CSI profile of the same room. Although the accuracy drops a little bit to 89% and 83% for location A when the density estimation is performed using the predefined CSI profiles of location B and C, respectively, our system still maintains a high accuracy. This results indicate that profiles built from the training rooms can be extended to other rooms since the pattern of the CSI distribution is preserved.

Different Number of People. In addition, we further explore the impact of profile construction from different number of participants. In particular, we use profiles built upon 3 people in each sub-region and use 4 during the testing phase. Table. 1 shows the EMD distance between the profile and the testing data for all four different regions in location A. Each cell contains the EMD distance value between the corresponding column (i.e., profile) and row (i.e., possible candidates). The gray cell indicates the minimum value among all the candidates and we observe that the minimum distance appears in the diagonal of the table. That means, our algorithm successfully predicts the people density region using profiles constructed from different numbers of people. The result also indicates the proposed system has the ability to extend the profile built from one location to another.

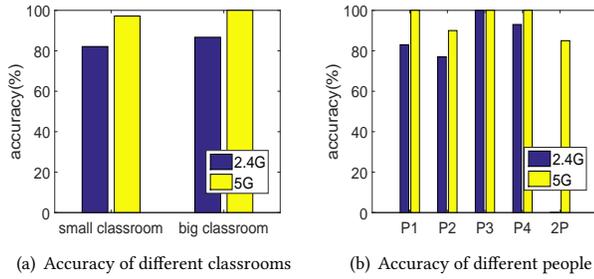


Figure 14: Performance of speed estimation utilizing THD-based algorithm

6.6 Walking Speed and Direction Derivation

We first examine the performance of walking speed estimation with the proposed THD-based method. As shown in Figure 14(a), the average accuracy of speed estimation at 5GHz outperforms that at 2.4GHz in both small classroom and big classroom due to the heavy interferences at 2.4GHz frequency band. In particular, our system achieves 96% and 99% accuracy for small classroom and big classroom at 5GHz whereas the accuracy for 2.4GHz are 82% and 85%. In addition, we also observe that the accuracy in the big classroom is better than in the small classroom due to less multi-path impacts.

Figure 14(b) shows the average speed estimation accuracy for different participants at 2.4GHz and 5GHz. We evaluate 4 participants which are represented as P1, P2, P3 and P4. The 98.6% average accuracy of overall performance on 5GHz across different participants is encouraging, but the estimation accuracy for participant 2 is 78% at 2.4GHz and 90% at 5GHz, which are lower than other participants. By checking the ground truth recorded by a camcorder, we find that the step frequency of participant 2 does not change while his moving speed increases. Therefore, the difference of the fundamental signal amplitudes with respect to different speeds becomes less significant and thereby results in the degradation on speed estimation. We also evaluate 2 participants walking as a group on 5GHz which is denoted as 2P in the figure. Even though the 85% percentage accuracy is slightly lower than the average accuracy of 1 person, two walking activities are still successfully captured by our algorithm, which makes our system more practical.

Next, we study the performance of walking direction estimation in our system. We show the walking direction classification results as a confusion matrix in Figure 15(a). The results show that our walking direction estimation method can achieve an average accuracy of 96.6% for the three walking directions. As shown in Figure 15(b), the direction estimation performance improves with the number of training samples. We can also observe that our direction estimation algorithm has better performance at 5G than 2.4G. This is because 5G suffers less interference over the wireless channel than 2.4G and thus provide more reliable CSI measurements.

7 CONCLUSION

The awareness of human dynamics information is increasingly important in designing management services, e.g., caring services

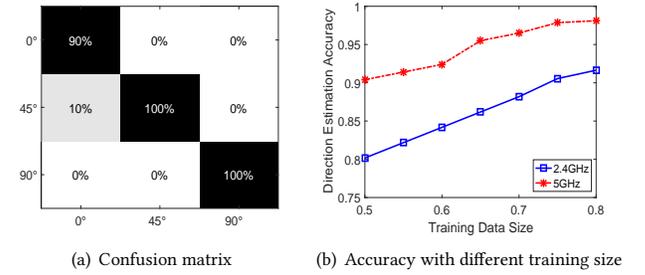


Figure 15: Performance of direction estimation using classification-based method

for seniors especially for those who have been identified as suffering from social isolation in assisted living places or smart facility management (e.g., understanding the popular activities or assisting evacuation planning) in meeting/activity rooms in campuses or corporate buildings. In this paper, we show that the prevalent WiFi signals can be exploited to provide a comprehensive view of the human dynamic information to facilitate social activeness assessment and smart environment design. Our approach is low-cost and device-free, which does not require any active human participation. In particular, our system provides fine-grained human dynamics monitoring via estimating the number of participants, human density and people walking speed and direction leveraging Channel State information (CSI). The semi-supervised learning approach based on non-linear regression can estimate the number of participants with significantly reduced training efforts. Thus, the constructed model can be generalized and applied to different environments, making the system easily scalable. Furthermore, the human density estimation grounded on the statistical distribution of CSI measurements provides complementary information to the number of participants to facilitate senior health services and smart space management. Additionally, our system develops a Total Harmonic Distortion (THD)-based mechanism along with a frequency-based classification approach to infer both walking speed and direction. Such detailed information on walking patterns is useful for scenarios where seniors need to be watched closely for their activities to ensure their safety and to provide in-time assistance. Our extensive experiments conducted in multiple types of indoor environments confirm the effectiveness of exploiting commodity WiFi to capture the human dynamic characteristics without attaching any monitoring devices to participants. The evaluation results demonstrate that our low-cost WiFi-based system is highly reliable and scalable on monitoring of human dynamics in indoor environments. We will explore the walking speed and direction estimation under the multi-people scenario in our future work.

8 ACKNOWLEDGMENTS

We thank our shepherd, Dr. Wen Hu and the anonymous reviewers for their insightful feedbacks. This work was partially supported by the National Science Foundation Grants CNS1514436, CNS1409767 and CSR1217379.

REFERENCES

- [1] Martin Azizyan, Ionut Constandache, and Romit Roy Choudhury. 2009. SurroundSense: mobile phone localization via ambience fingerprinting. In *Proceedings of the ACM International Conference on Mobile Computing and Networking (ACM MOBICOM)*. 261–272.
- [2] Siu-Yeung Cho, Tommy WS Chow, and Chi-Tat Leung. 1999. A neural-based crowd estimation by hybrid global learning algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 29, 4 (1999), 535–541.
- [3] Saandeep Depatla, Arjun Muralidharan, and Yasamin Mostofi. 2015. Occupancy estimation using only wifi power measurements. *IEEE Journal on Selected Areas in Communications* 33, 7 (2015), 1381–1393.
- [4] Robyn A Findlay. 2003. Interventions to reduce social isolation amongst older people: where is the evidence? *Ageing and Society* 23, 5 (2003), 647–658.
- [5] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review* 41, 1 (2011), 53–53.
- [6] Chunmei Han, Kaishun Wu, Yuxi Wang, and Lionel M Ni. 2014. WiFall: Device-free fall detection by wireless networks. In *Proceedings of the IEEE International Conference on Computer Communications (IEEE INFOCOM)*. 271–279.
- [7] <https://www.helpguide.org/articles/depression/depression-in-older-adults-and-the-elderly.htm>. 2017. Depression in Older Adults. (2017).
- [8] <https://www.nia.nih.gov/health/publication/participating-activities-you-enjoy>. 2015. Participating in Activities You Enjoy. (2015).
- [9] <https://www.osha.gov/SLTC/etools/evacuation/evac.html>. 2017. Emergency Action Plan. (2017).
- [10] Pravein Govindan Kannan, Seshadri Padmanabha Venkatagiri, Mun Choon Chan, Akhihebbal L Ananda, and Li-Shiuan Peh. 2012. Low cost crowd counting using audio tones. In *Proceedings of the ACM Conference on Embedded Network Sensor Systems (ACM Sensys)*. 155–168.
- [11] Zhenjiang Li, Yaxiong Xie, Mo Li, and Kyle Jamieson. 2015. Recitation: Rehearsing wireless packet reception in software. In *Proceedings of the ACM International Conference on Mobile Computing and Networking (ACM MOBICOM)*. 291–303.
- [12] Sifeng Liu and Yi Lin. 2005. *Grey Information: Theory and Practical Applications (Advanced Information and Knowledge Processing)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- [13] Ruihua Ma, Liyuan Li, Weimin Huang, and Qi Tian. 2004. On pixel count based crowd density estimation for visual surveillance. In *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems (IEEE CIS)*. 170–173.
- [14] M Nakatsuka, H Iwatani, and J Katto. 2008. A study on passive crowd density estimation using wireless sensors. In *Proceeding of the International Conference on Mobile Computing and Ubiquitous Networking (Citeseer ICMU)*.
- [15] Yossi Rubner and Carlo Tomasi. 2013. *Perceptual metrics for image database navigation*. Vol. 594. Springer Science & Business Media.
- [16] Lorenz Schauer, Martin Werner, and Philipp Marcus. 2014. Estimating crowd densities and pedestrian flows using wi-fi and bluetooth. In *Proceedings of the International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (ICST MOBIQUITOUS)*. 171–177.
- [17] Michael Von Korff, Susan M Shortreed, Linda LeResche, Kathleen Saunders, Stephen Thielke, Manu Thakral, Dori Rosenberg, and Judith A Turner. 2017. A Longitudinal Study of Depression among Middle-aged and Senior Patients Initiating Chronic Opioid Therapy. *Journal of Affective Disorders* 211 (2017), 136–143.
- [18] Wei Wang, Alex X Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*. 363–373.
- [19] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. 2014. E-eyes: device-free location-oriented activity identification using fine-grained WiFi signatures. In *Proceedings of the ACM International Conference on Mobile Computing and Networking (ACM MobiCom)*. 617–628.
- [20] Jens Weppner, Benjamin Bischke, and Paul Lukowicz. 2016. Monitoring crowd condition in public spaces by tracking mobile consumer devices with wifi interface. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM UbiComp)*. 1363–1371.
- [21] Jens Weppner and Paul Lukowicz. 2011. Collaborative crowd density estimation with mobile phones. In *Proceedings of ACM PhoneSense (Citeseer)*.
- [22] Chenshu Wu, Zheng Yang, Zimu Zhou, Kun Qian, Yunhao Liu, and Mingyan Liu. 2015. PhaseU: Real-time LOS identification with WiFi. In *Proceeding of the IEEE Conference on Computer Communications (IEEE INFOCOM)*. 2038–2046.
- [23] Dan Wu, Daqing Zhang, Chenren Xu, Yasha Wang, and Hao Wang. 2016. WiDir: walking direction estimation using wireless signals. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (ACM Ubicomp)*. 351–362.
- [24] Wei Xi, Jizhong Zhao, Xiang-Yang Li, Kun Zhao, Shaojie Tang, Xue Liu, and Zhiping Jiang. 2014. Electronic Frog Eye: Counting Crowd Using WiFi. In *Proceedings of the International Conference on Computer Communications (IEEE INFOCOM)*. 361–369.
- [25] Chenren Xu, Bernhard Firner, Robert S Moore, Yanyong Zhang, Wade Trappe, Richard Howard, Feixiong Zhang, and Ning An. 2013. SCPL: indoor device-free multi-subject counting and localization using radio signal strength. In *Proceeding of the ACM/IEEE International Conference on Information Processing in Sensor Networks (IEEE IPSN)*. 79–90.
- [26] Yunze Zeng, Parth H Pathak, and Prasant Mohapatra. 2016. WiWho: WiFi-based Person Identification in Smart Spaces. In *Proceeding of the ACM/IEEE International Conference on Information Processing in Sensor Networks (IEEE IPSN)*. 1–12.
- [27] Shaghayegh Zihajehzadeh and Edward J Park. 2016. Experimental evaluation of regression model-based walking speed estimation using lower body-mounted IMU. In *Proceedings of the IEEE International Conference of the Engineering in Medicine and Biology Society (IEEE EMBC)*. 243–246.