# Human Detection For Crowd Count Estimation Using CSI of WiFi Signals

Omotayo Oshiga, Hussein U. Suleiman, Sadiq Thomas, Petrus Nzerem, †Labaran Farouk and Steve Adeshina
Department of Electrical & Electronics Engineering, Nile University of Nigeria, Abuja, Nigeria;
†Airforce Institute of Technology, Kaduna
Email: [ooshiga;h.suleiman;sadiqthomas;petrus.nzerem;steve.adeshina]@nileuniversity.edu.ng
Email: †ismail.labaran@airforce.mil.ng

***Abstract*—We address the problem of crowd estimation in situations such as indoor events using anonymous and non-participatory CSI of WiFi Signals. Observing the great resemblance of Channel State Information (CSI, a finegrained information captured from the received Wi-Fi signal) to texture, we propose a brand-new framework based on statistical mechanics, and relying only on sets of machine learning techniques.**

**In this paper, a framework for crowd count estimation is presented which utilizes Chebyshev filter and SVD to remove background noise in the CSI data, PCA to reduce the dimensionality of the CSI data and spectral descriptors for feature extraction. From the extracted feature, a set of classiffying algorithms are then utilised for training and testing the accuracy of our crowd estimation framework The aim of this framework to effectively and efficiently extract the channel information in WiFi signals across OFDM carriers reflected by the presence of human bodies. From the experiments conducted, we demonstrate the feasibility and efficacy of the proposed framework. Our result depict that our estimation becomes more–rather than less–accurate when the crowd count increases.**

## I. Introduction

WiFi signals has caught tremedous attentions for it's low cost, privacy-preserving and ubiquity. Motivated by these, researchers have well investigated into its potential for human detection and localization. Therefore, the question is how much information do WiFi carry about humans? For instance, can a WiFi link count the number of people present in an area without them carrying any device? Can this be done with WiFi channel variations?

A survey of the literature indicates that the problem of crowd counting is currently been investigated by researchers from computer-vision, wireless networking, and environmental science communities [1]–[5], [7]–[9]. In this paper, a framework for crowd estimation is presented which utilizes Chebyshev filter and SVDto remove background noise in the CSI data, PCA to reduce the dimensionality of the CSI data and spectral descriptors for feature extraction. From the extracted feature, a set of classiffying algorithms are then utilised for training and testing the accuracy of our crowd estimation framework The aim of this framework to effectively and efficiently extract the channel information in WiFi signals across OFDM carriers reflected by the presence of human bodies. From the experiments conducted, we demonstrate the feasibility and efficacy of the proposed framework. Our result depict that our estimation becomes more–rather than less–accurate when the crowd count increases. Therein, our contributions include: We provide an overview on OFDM channel state information and MIMO in Wi-Fi. We present an Human Dection/Crowd Count Estimation Framework using three main steps - CSI Pre-processing, Human Detection and Multiple Antenna Detection to enhance the accuracy of our crowd estimation model. We further experimentally verify the possibility of using our crowd estimation framework to count people, and discuss some factors that may affect performance.

## II. Preliminaries

### A. Channel State Information : Background and Formulation

Currently, many modern wireless communication systems and standards include WPAN interfaces such as IEEE 802.11a/g/n, ultra-wideband (UWB) and 4G Networks such as LTE. In these systems, OFDM (Orthogonal Frequency Division Multiplexing) is implemented so as to reduce interference, cut-out and fading due to multipath propagation. OFDM is a digital modulation scheme which segments the bandwidth into several closely-spaced orthogonal frequencies called subcarriers, with each carrying a data stream. In current OFDM systems as shown in Figure 1, the CSI is a fine-grained information from the physical layer that describes signal propagation – Channel Frequency Response (CFR) – from the transmitter to the receiver at the subcarrier level. It can reveal the effects of fading and power decay with distance based on the amplitude and phase response of each subcarrier within a channel. Therefore, the presence of a human body imprints on intercepted multipath components and hence on the CSI, allowing the recognition of its presence.

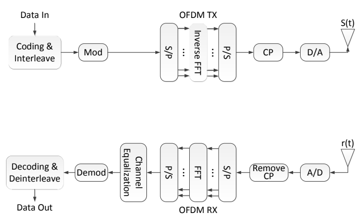


**Figure 1:** The Framework of Orthogonal Frequency Division Multiplexing (OFDM)

The OFDM-based transmission encoding data stream on a set of orthogonal subcarriers is modelled as:

$$Y = \mathbf{H} \cdot X + N, \tag{1}$$

where $Y$ and $X$ are the receive and transmit complex vectors, $\mathbf{H}$ is the complex channel frequency response or CSI matrix, which denotes the overall influence of the environment to $X$ and $N$ is the noise vector. After decoding the received signal, the receiver can learn the CSI $H$ by comparing $Y$ and $X$. The proposed crowd count estimation system seek to exploit how human presence causes time variations in the channel frequency response of the CSI. The raw complex CSI data to be collected contains the information of $N$ subcarriers conveyed as

$$H = \Big[h(f_1), h(f_2), \cdots, h(f_N)\Big]^{\mathrm{T}}, \tag{2}$$

where $h(f_n)$ is the CSI at subcarrier $n$ with transmission frequency $f_n$. The $n$-th subcarrier in $H_{ij}$ is expressed as

$$h(f_n) = \big|h(f_n)\big| e^{j\angle h(f_n)}, \;\; n \in [1, N], \tag{3}$$

where $\big|h(f_n)\big|$ is the amplitude response and $\angle h(f_n)$ is the phase response of the subcarrier $n$. It is important to note that the Signal-to-Noise Rstio is used to represent amplitude using $SNR_n = 10\log_{10}(\big|h(f_n)\big|^2)$ and seperate subcarriers experience different multipath fading due to different transmission frequency $f_n$.

### B. Multiple Input and Multiple Output (MIMO)

The use of multiple antennas technology at both the transmitter and receiver has transformed wireless and mobile communication in the last decade. These antennas at each end of the communication spectrum are combined to minimize errors and optimize data transfer. The use of multiple antennas at each end of the communication spectrum, eliminates the trouble caused by multipath propagation, and can sometimes even take advantage of this effect. Multiple Input and Multiple Output (MIMO) is one of the several forms of smart antenna technology others being Multiple Input and Single Output (MISO) and Single Input and Multiple Output (SIMO). Inherently in WLAN, the MIMO technique is the key feature of the IEEE 802.11n which sets it apart from the earlier IEEE 802.11a/g versions.

MIMO as a smart technology uses multiple antennas to maximize signal power via multi-stream beamforming, increase reliability via spatial diversity and improve data throughput via spatial multiplexing. In MIMO, each combination of transmitting (TX) and receiving (RX) antenna is considered to be a seperate data link (TX-RX link). Suppose there are $N_{tx}$ transmitting antennas and $N_{rx}$ receiving antennas, the CSI stream $\mathbf{H}$ is a $(N_{tx} \times N_{rx} \times N)$ matrix to denote the group of CSIs of all TX-RX links can be expressed as:

$$\mathbf{H} = (H_{ij})_{N_{tx} \times N_{rx}} = \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1N_{rx}} \\ H_{21} & H_{22} & \cdots & H_{2N_{rx}} \\ \vdots & \cdots & \vdots & \\ H_{N_{tx}1} & H_{N_{tx}2} & \cdots & H_{N_{tx}N_{rx}} \end{bmatrix}, \tag{4}$$

where $H_{ij}$ is the CSI vector containing the information of all $N$ subcarriers formed by TX $i$ and RX $j$, which is

$$H_{ij} = \Big[h(f_1), h(f_2), \cdots, h(f_N)\Big]^{\mathrm{T}}, i \in [1, N_{tx}], j \in [1, N_{rx}]. \tag{5}$$

Currently, the most widely used tool for CSI extraction is the Intel CSI tool in [1]. For the purpose of our experiments, we adopt this Intel 5300 NIC CSI tool and focus our investigation on the CSI amplitude.

Due to the use of multiple subcarriers, separate CSI streams experience different multipath fading. When small movements occur in the environment, individual subcarrier measurements are likely to change. Therefore, CSI portrays a fine-grained spectral structure of wireless links and exhibits properties suitable for human presence detection. As seen in Figure 2, CSI exhibits different patterns across subcarriers when there exist human absence or presence in the area of interest. This illustrates the properties of a CSI stream for given TX-RX link for device-free presence detection. For a chosen set of subcarriers, it is seen that the packet-axis represents the packets over time and the SNR-axis represents the amplitude of each subcarrier. The different amplitude levels in the area of interest illustrates that human presence can be detected through classification on CSI patterns.

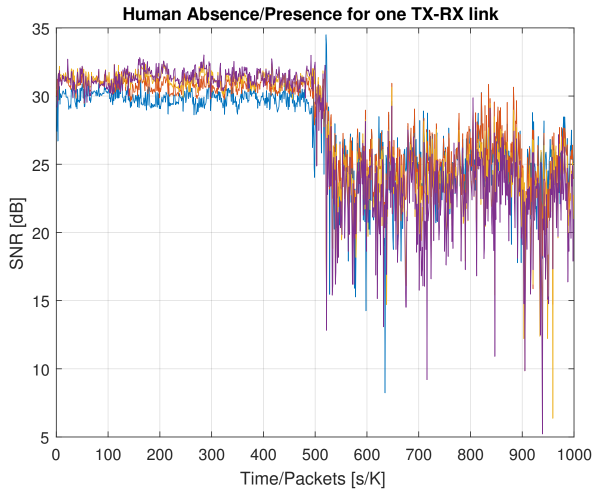


**Figure 2:** CSI amplitude at subcarrier 1, 10, 20, 30 over time/packets for given a TX-RX link.

## III. Proposed Crowd Count Estimation Framework

In this section, we present the flow of the overall flow of the proposed framework as seen in Figure 3. This framework initializes by propagating WiFi signals from a commercial wireless router acting as a transmitter (TX) and performing CSI collection using the commercial Intel 5300 NIC acting as a receiver (RX) on a laptop computer. With $N_{tx}$ transmitting antennas and $N_{rx}$ receiving antennas, we have $M = N_{tx} \times N_{rx}$ TX-RX pairs, each generating a set of CSI stream with dimensions $N = 30$ (subcarriers) $\times$ $K$ (Number of samples). Next, we then process all the sets of CSI stream seperately and investigate their performance based on early or late fusion of the CSI streams.

Since, the raw CSI streams contain noise and interference caused by other devices in the WiFi channel, we first pass them through a data pre-processing scheme. The CSI pre-processing scheme include an outlier removal method using OPTICS with outlier factors, CSI replacement using linear interpolation, noise removal using a Chebyshev filter and a location dependency removal using singular value decomposition (SVD) [12], [13], [17], [18]. Although this process helps in removing high frequency noises, the noises in the filtered CSI signal are not completely eliminated, which are mainly caused by internal state changes in the . Though, a strict low-pass filter can further remove all the noise in the signal, but will lead to significant loss of useful information. To further filter the CSI data, we apply PCA-based filtering which retains the necessary contributing information to human detection/crowd estimation while futher reducing the correlation between different subcarriers and the dimensionality of each CSI stream. As it is known the presence of humans causes changes in the time variations of a CSI time series leading to changes in the spread of the subcarriers. Hence, we propose to extract spectral descriptors which quantize the time variations in the CSI stream as robust features for human detection/crowd estimation. The extracted features are then fed to a modern classification model to determine human absence or presence. With human presence, they are further utilized to determine the number of humans present. Finally, utilizing the multiple antennas technique in MIMO,the extracted features of every TX-RX pairs are either early fused together before been fed to the machine learning classification model to obtain a result or late fused together by applying a majority vote on the results of every TX-RX pairs.

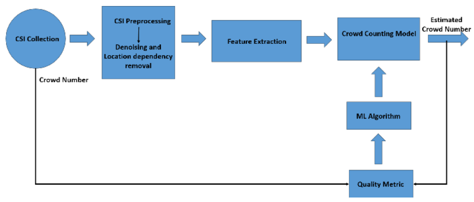


**Figure 3:** The proposed framework for Human Detection/Crowd Count Estimation

### *A. CSI Pre-processing*

Analysing the sequential packets of a CSI data with each packet having $M$ pairs of TX-RX links, observations show that each human has specific influence on WiFi signals due to the difference in body shapes and motion patterns as seen in Figure 2 for one of the TX-RX links. Hence for all the closely spaced subcarriers, it is clear that variations in the CSI time series due to human movements around the transmitter and receivers are correlated and noisy for each TX-RX links. It is important to note that the noise in the raw CSI data result from other WiFi devices and are location-dependent (contain information about a specific environment), therefore a Chebyshev filter and singular value decomposition (SVD) would be investigated to remove background noise in the CSI data related to a specific environment. The main idea of this action is to ensure that the training data and features to be later extracted from the CSI data are similar for different indoor environments. As a result, we elaborate on the methodology required to perform CSI pre-processing to clean and interpolate the CSI data of each TX-RX link for each sequential packet.

*1) Outlier Removal:* : The first step to process CSI is to remove outliers. During our analysis of CSI data, we observed that there existed anomalous measurements which were not caused human presence and needed to be eliminated before human detection. Figure 2 illustrates the outliers contained in the raw CSI data collected in the absence of human. To identify and remove these outliers, we apply the OPTICS: Ordering Points To Identify the Clustering Structure on all 30 subcarriers.

*2) Linear Interpolation:* : It should be noted that although we have set the transmitter to transmit 50 packets per second in our experiments, we cannot guarantee the sampling rate of CSI with the same frequency. In fact, we find that sampling jitter is quite common in CSI, resulting from packet loss and some other reasons. Therefore, to account for the CSI information loss caused by sampling jitter and outlier removal, the CSI measurements must be interpolated. Specifically, we utilize the 1-D linear interpolation algorithm to ensure evenly spaced CSI with 20ms apart between consecutive measurements.

*3) Noise Removal:* First, due to interference caused by commercial WiFi devices in the same channel, packets received are extremely noisy and not evenly distributed in time. Since, the power distributions of different subcarriers vary, normalizing each subcarrier by subtracting the average of a moving-window, from each sample was performed. It is important to note that the frequency variations caused due to the movements of human usually lie at the low end of the spectrum while the frequency of the noise lies at the high end of the spectrum. Thus, We linear interpolated of the raw CSI data to 100 samples per second. the, to remove the high frequency noise, a 5th-order Chebyshev Type 2 low-pass filter with a cut-off frequency of 50Hz is applied (on on each subcarrier for different transmission links) as it does not distort the phase information in the signal and has a maximally flat amplitude response in the passband while maintain human motion signals (variations in the CSI time series) as seen in Figure 4. Though this process helps in removing high frequency noise, the noise in the signal is not completely eliminated. Although, a strict low-pass filter can further remove all the noise in the signal, but will lead to significant loss of useful information.

*4) Location Dependency Removal:* To deal with location dependency, a method based on SVD to remove the location-dependent background energy would be utilized. The CSI matrix $\mathbf{H}$ of each TX-RX link has dimension $K \times N$ (K is total number of, after applying SVD, $\mathbf{H}$ is decomposed into three matrices, $\mathbf{U}$, $\mathbf{S}$ and $\mathbf{V}$, where $\mathbf{U}$ is a $K \times K$

orthogonal matrix, $\mathbf{V}$ is a $N \times N$ orthogonal matrix and $\mathbf{S}$ is a $K \times N$ matrix with diagonal elements representing the singular values of $\mathbf{H}$. After decomposition, the location-dependent background energy is removed by replacing the first singular value $SV_1$ with zero. If we set the all SV in $\mathbf{S}$ to zero except for the largest one, denoted by $\mathbf{S} = \mathbf{S}_{SV_1=0}$ and reconstruct the channel state information matrix using $\mathbf{H}_{SV_1=0} = \mathbf{U}\mathbf{S}_{SV_1=0}\mathbf{V}^{\mathrm{T}}$, then the reconstructed CSI $\mathbf{H}_{SV_1=0}$ would be similar to the CSI recorded in a quite environment without background energy or noise. This shows that the largest SV packs most of the background energy corresponding to static paths. Therefore, the largest SV is set to zero so that the background energy can be mostly be removed and the reconstructed CSI preserves human actions. For the CSI data of different environments, after removing the background energy (largest SV), the reconstructed CSI data should be similar to each other.

### B. Human Detection

*1) PCA-Based Dimensionality Reduction:* After CSI denoising and removal of background energy in the CSI data using Chebyshev filter and singular value decomposition (SVD) to remove background noise in the CSI data related to a specific environment. The sequential packets of CSI data collected over a period of time tis then prepared for feature extraction. It is important to remember that a CSI stream $H_{ij}$ formed by TX $i$ and RX $j$ has $N = 30$ subcarriers, therein 30 dimensions of high complexity leads to time consumption. Since each dimension may have different contribution to the presence detection, we apply Principal Component Analysis (PCA) to extract the most contributing features and reduce the dimensionality of each CSI stream. PCA converts a set of possibly correlated variables into a set of linearly uncorrelated variables, called principal components, through orthogonal transformation. The first principal component has the largest possible variance, and each succeeding component in turn has the highest variance possible and orthogonal to the preceding components. For the goal of dimensionality reduction, PCA is to find $L = 5$ new dimensions (principal components), with each one being a linear combination of the original dimensions, so that the new dimensions can reveal the nature of the original data and compress them. Therefore, using PCA, a single CSI stream $H_{ij}$ for each TX$i$ -RX$j$ can be reduced from $N = 30$ subcarriers to $L = 5$ principal components in Figure 5, where the reduced CSI stream $\tilde{H}_{ij}$ formed by TX $i$ and RX $j$ is denoted as

$$\tilde{H}_{ij} = \left(\tilde{h}_1, \tilde{h}_2, \cdots, \tilde{h}_L\right)^{\mathrm{T}}, \;\; i \in [1, N_{tx}], \;\; j \in [1, N_{rx}]. \quad (6)$$

*2) Feature Extraction:* For feature extraction, every $m = 10$ sequential packets of the reduced CSI streams $\tilde{H}_{ij}$ obtained within time $t = 5$secs window are utilized to perform feature extraction for human presence detection. As such, feature vectors are to be extracted from each TX$i$ -RX$j$ pair of $m$ sequential packets of reduced CSI streams $\tilde{H}_{ij}$. These $m$ sequential packets of reduced CSI streams are denoted as $\mathbb{H}_{ij} = \left[(\tilde{H}_{ij})_1, (\tilde{H}_{ij})_2, (\tilde{H}_{ij})_3, \cdots, (\tilde{H}_{ij})_m\right]$.

As it is known the presence of humans causes changes in the time variations of a CSI time series leading to changes in the spread of the sub-carriers. Here, effective spectral descriptors would be utilized to quantize the time variations in $\mathbb{H}_{ij}$. These spectral descriptors $\{F_d\}_{d=1}^{D=12}$ include the mean, standard deviation, centroid, spread, 1st and 2nd-order moment, mean-sigma ratio, number of peaks, number of valleys, maximum width between peaks, maximum width between valleys, and width/height between the maximum peak and minimum valley. Therefore, new features $\mathbf{F}_k$ (dimension $D \times L$) are extracted for each TX$i$ -RX$j$ link by applying these spectral descriptors $F_d$ on $\mathbb{H}_{ij}$. Therein, for each of the spectral descriptors $F_d$ , $L$ new features are to be obtained from $\mathbb{H}_{ij}$. Therefore, the new feature vectors $\mathbf{F}_k$ for each $M$ TX$i$ -RX$j$ pairs are denoted as

$$\mathbf{F}_k = \left[F_1(\mathbb{H}_{ij}), F_2(\mathbb{H}_{ij}), F_3(\mathbb{H}_{ij}), \cdots, F_D(\mathbb{H}_{ij})\right], \quad (7)$$

where $F_d(\mathbb{H}_{ij}) = \left[f_1^d, f_2^d, \cdots, f_L^d)\right]$ is any given spectral descriptor.

*3) Crowd Estimation/Presence Detection by a Training Classifier:* Classification consists of training and testing. Each training sample contains a class label and several features. Classification is to determine the labels of the testing samples according to their features, based on the models established using the training samples. For the problem of presence detection, two kinds of CSI samples need to be collected during training: (1) one (1) labelled samples collected with human presence; (2) zero (0) samples collected with human absence. Assume $K$ is the number of training samples, $D \times L$ denotes the dimensionality of the features. Each training sample consists of $M$ TX-RX links $(\mathbf{F}_k, C_k)$, where $\mathbf{F}_k$ is the feature vector and $C_k \in \{0, 1\}$ is the label assigned during training, with 1 representing human presence and 0 representing human absence. These labeled samples are used to establish the training classifiers. Assume $(\mathbf{F}_1, C_1)$ is a testing sample, with $\mathbf{F}_1 \in \mathcal{R}^{D \times L}$ as the CSI stream for each $M = 6$ TX-RX links, classification is to determine the value of $C$, i.e. to determine human presence or absence in the area of interest.

Next, we investigate these new features by fusing the feature vectors of all TX-RX links before or after training the classifiers. 1) Early Fusion: We concatenate the features from $M$ TX-RX links into a new feature. Then, we train a single (1) classifier and take action with the highest probability as the predicted result. 2) Late Fusion: Instead of concatenating six feature vectors and training a single classifier, we train one for each pair, so there would be six (6) train classifiers. Given a testing instance, a probability of each detection is obtained, rendering six probability vectors of length detection (two – human presence or absence). Summing these six vectors, we take the detection with the highest probability as the predicted result. For the experiments 3 classifing algorithms were utilized for our analysis using various settings, in which


**Figure 4:** Computer Networking Laboratory where the experiment was performed

yCLF means training y CLF classifiers – Support Vector Machine (SVM), Random Forest (RF) and Gradient Boosting Classifiers (GB).

### C. Multiple-Antenna Enhancement

With the endorsement of MIMO in modern communication, more and more WiFi devices are equipped with multiple antennas. Since not all wireless links are equally sensitive to human movement and the sensitivity varies with link fade level along with other factors, we propose a majority-vote based detection algorithm to combat existing bad stream during CSI collection. To be more specific, we apply human detection on each TX-RX antenna pair respectively and fuse all the detection results with a majority-vote scheme for more precise and reliable detection.

## IV. Experimental Evaluation

In this section, we will confirm the effectiveness of the proposed techniques through experimental evaluations. In order to verify the proposed framework, we performed the experiment in the computer networking laboratory at Nile University as shown in Figure 4.

The infrastructure of this experiment is composed of a TL-WDR4300 wireless router supporting IEEE 802.11n (OFDM and MIMO) for signal transmission and a computer for CSI extraction and processing. The computer is installed with commercial off-the-shelf Intel 5300 NICs supporting IEEE 802.11n (OFDM and MIMO) operating at 2.4GHz with modified firmware to retrieve raw complex CSI values for $N = 30$ subcarriers using the tool released in [1]. With $N_{tx} = 3$ transmitting antennas and $N_{rx} = 2$ receiving antennas, we have $M = 6$ TX-RX pairs, each generating a set of CSI stream with dimensions $N \times K$ . These actions are performed in a 5-second period, each generating $M = 6$ six sets of CSI with dimension $N \times K = 30 \times 5000$ (interpolated to 1000 samples/second). We then process the CSIs separately and investigate whether early or late fusion yields better performance.

These CSI Data are collected for verifying if our proposed method could accurately detect human presence for determining the number of people in a room. Therein, we invite 10 people - five male and five female subjects and sub-divide them into groups of $\{0, 1, 2-3, 4-6, 7-10\}$ people in the room and ask them to randomly walk into the seminar room.

For each groups, we evaluate the performance of each classifier (SVM, RF and GB – early and late fusion) with the proposed framework – Chebyshev filter and SVD to remove background noise in the CSI data, PCA to reduce the dimensionality of the CSI data and spectral descriptors for feature extraction using 10-fold cross validation.

Results in Table I shows the results of crowd estimtation for different number of people in a room using early fusion and late fusion. As the statistics show, late fusion performs better since it exploits six different channels with six different classifiers. Though each classifier is weaker compared to that of early fusion, more channels provide more information for human presence detection. In addtion, the results among all groups of people in the room show tendency that when more training data are involved, the higher accuracy may be achieved, implying the model is improved by investigating more data.

## V. Conclusions

In this paper, we proposed a framework for crowd estimation utilizing Chebyshev filter and SVD to remove background noise in the CSI data, PCA to reduce the dimensionality of the CSI data and spectral descriptors for feature extraction to effectively and efficiently extract the channel information of WiFi signals reflected by the presence of human bodies. From the extracted feature, a set of classiffying algorithms are then utilised for training and testing the accuracy of our crowd estimation framework

From the experiments conducted, we demonstrated the feasibility and efficacy of the proposed methods. Finally, our result depicted that our estimation becomes more–rather than less–accurate when the crowd count increases. Our experimental results show that this solution achieves an accuracy above 95% classifying the five groups of crowd number.

## References


[1] D. Halperin, W. Huy, A. Shethz, and D. Wetherall,"Tool release: Gathering 802.11n traces with channel state information," ACM SIGCOMM Computer Communication Review , January 2011.

[2] Z. Yang, Z. Zimu, and Y. Liu,"From RSSI to CSI: Indoor localization via channel response," ACM Computing Surveys (CSUR) 46.2 (2013): 25.

[3] S. Depatla, A. Muralidharan, and Y. Mostof,"Occupancy estimation using only WiFi power measurements," IEEE Journal on Selected Areas in Communications 33, no. 7 (2015): 1381-1393.

[4] Y. Yuan, Z. Jizhong, C. Qiu, and W. Xi,"Estimating crowd density in an RF-based dynamic environment," IEEE Sensors Journal 13, no. 10 (2013): 3837-3845.


**TABLE I:** Accuracy of Early and Late Fusion of the Training Classifier.

| Classifier/Count | 0 | 1 | 2-3 | 4-6 | 7-10 |
|---|---|---|---|---|---|
| 1SVM | %99.17 | %97.50 | % 97.50 | % 96.67 | % 96.67 |
| 6SVM | %98.33 | %96.67 | %99.17 | % 96.67 | % 96.67 |
| 1RF | %99.17 | %97.50 | % 97.50 | % 96.67 | % 96.67 |
| 6RF | %98.33 | %96.67 | %99.17 | % 96.67 | % 96.67 |
| 1GB | %99.17 | %98.33 | %99.17 | % 99.17 | % 99.17 |
| 6GB | %99.17 | %99.17 | % 98.33 | % 99.17 | %100 |

[5] M. Nakatsuka, H. Iwatani, and J. Katto,“A study on passive crowd density estimation using wireless sensors,” 4th Intl. Conf. on Mobile Computing and Ubiquitous Networking, 2008.
[6] T. Yoshida, and Y. Taniguchi,“Estimating the number of people using existing wifi access point in indoor environment,” In Proceedings of the 6th European Conference of Computer Science, pp. 46-53, 2015.
[7] S. H. Doong, “Spectral human flow counting with rssi in wireless sensor networks,” IEEE International Conference on Distributed Computing in Sensor Systems, pp. 110-112, 2016.
[8] W. Xi, Z. Jizhong, et al,“Electronic frog eye: Counting crowd using wifi,” IEEE INFOCOM, pp. 361-369, 2014.
[9] D. Domenico,G. Pecoraro, E. Cianea, and M. De Sanctis,“Trained device-free crowd counting and occupancy estimation using WiFi: A Doppler spectrum based approach,” IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications, pp. 1-8, 2016.
[10] D. Domenico, E. Cianea, and M. De Sanctis,“Trained device-free crowd counting and occupancy estimation using differential wifi channel state information,” Proceedings of the 3rd International on Workshop on Physical Analytics. ACM, 2016.
[11] y. LeCun, Y. Bengio, and G. Hinton, “Deep learning.” Nature 521, no. 7553 (2015): 436-444.
[12] G. E. Hinton, and R. R. Salakhutdinov. “Reducing the dimensionality of data with neural networks.” science 313, no. 5786 (2006): 504-507.
[13] V. Vapnik,“The Nature of Statistical Learning Theory,” Springer, New York, 1995.
[14] D. Halperin, W. Huy, A. Shethz, and D. Wetherall,“Predictable 802.11 packet delivery from wireless channel measurements,” ACM SIGCOMM Computer Communication Review , Vol. 40. No. 4, pp. 159-170, 2010.
[15] Sebastian Ruder,“Transfer Learning - Machine Learning’s Next Frontier ,” March 21, 2017 - Online [http://ruder.io/transfer-learning/index.html#whatistransferlearning].
[16] A Pan, Sinno Jialin, and Qiang Yang. ”A survey on transfer learning.” IEEE Transactions on knowledge and data engineering 22, no. 10 (2010): 1345-1359.
[17] M. M. Abdalla, M. B. Abuitbel, and M. A. Hassan, “Performance evaluation of direction of arrival estimation using MUSIC and ESPRIT algorithms for mobile communication systems,” in *6th Joint IFIP Wireless and Mobile Networking Conference (WMNC’13)*, 2013, pp. 1–7.
[18] M. Ankerst, M. M. Breunig, H. Kriegel and J. Sander, ”OPTICS: Ordering Points To Identify the Clustering Structure,” ACM Press, 49–60, 1999.