WiHACS: Leveraging WiFi for Human Activity Classification using OFDM Subcarriers' correlation

Tahmid Z. Chowdhury*, Cyril Leung*, Chun Yan Miao†
*School of Electrical & Computer Engineering, The University of British Columbia
†School of Computer Science & Engineering, Nanyang Technological University

***Abstract*—Human Activity Classification using Wi-Fi Channel characteristics serve a diverse range of applications such as monitoring hospital patients and promoting independent living of elderly people. Current systems' key limitation is performance degradation when detecting activities across multiple walls or under non-line-of-sight (NLOS). Although their performance can be improved by engaging multiple transmitters/receivers, it creates an overhead in practical scenarios. We propose WiHACS: A *Wi-Fi* based *Human Activity Classification* based on OFDM *Subcarriers'* correlation, which engages a single AP and laptop. We employ correlation patterns across a particular range of subcarriers as our salient feature to detect human activities in different environments. Preliminary evaluations show WiHACS can correctly classify human activities with an average accuracy of 97%, 92%, and 75% in LOS, across one-wall and across two-walls respectively.**

***Keywords*—Activity Recognition, CSI, WiFi, Phase**

I. Introduction

Classifying human activities is an emerging field of research in pervasive computing due to its application in human-centric problems such as monitoring patients in hospitals, intrusion detection, and battlefield military applications [1]. Due to advent in sensor technologies and computer vision, most human activity classification systems categorize into two types: wearable and vision-based. Wearable devices require the subject/s to carry devices in their body such as smart-phones and smart-watches with embedded sensors and a survey on this can be found in [2]. Vision-based systems utilize computer vision and image processing techniques to classify activities; a comprehensive survey is available in [3]. The biggest drawback of sensor and vision based systems are, the former requiring subject to wear the device at all times and the latter invading privacy.

To overcome these issues researchers foresee the potential of utilizing wireless signals to detect human activities. This is because major developments in wireless communications and mobile computing have enabled the rise of IoT for smart-home applications [4]. Current Wi-Fi based systems categorize into-

(i) Received Signal Strength Indicator (RSSI) based: These systems measure RSSI in wireless links and map changes in this quantity to human activities. In [5] a model was proposed to localize and detect a moving target by utilizing RSSI variance of small-scale fading. A kernel distance-based radio tomographic was proposed in [6] to localize human beings. By utilizing software radio to improve RSSI granularity Sigg *et al.* improved classification of 4 human activities upto 72% [7]. Due to RF interference and wireless signals suffering from multipath fading, these systems degrade in performance due to the RSSI at a transmitter-receiver link being unreliable.

(ii) Specialized Radio Hardwares: These require Universal Software Radio Peripheral (USRP) or Software Defined Radio (SDR) to process PHY layer information. Pu *et al.* proposed WiSee [8] that classifies nine human body gestures by measuring Doppler shifts through a USRP with an accuracy of 95%. Kellogg *et al.* implemented a specialized analog circuit to utilize the envelope of amplitude of recieved signals to recognize gestures within 2.5 feet [9]. The drawback of such technologies is these are not commercial-off-the-shelf (COTS) devices and is expensive.

(iii) Channel State Information (CSI): This physical layer information can also be leveraged using commercial-off-the-shelf (COTS) Wi-Fi devices in the granularity of the OFDM subcarriers by modifying Linux drivers for an Intel 5300 [16] Network Interface Card (NIC). Zhang *et al.* utilized the CSI phase information to detect human falls [14]. Xi *et al.* [10] presented an algorithm to estimate carrier frequency offset (CFO) in WiFI devices and used the CSI phase variance to detect human activities. Zhu *et al.* [11] leveraged the CSI and the angle of arrival (AoA) to localize and detect a human activity.

Limitations of Existing Systems

Current systems utilizing CSI either localize a target [11], detect human gestures [12], classify a particular type of activity, such as falls of an elderly [14] or detect human activities [15]. But there exists performance degradation when classifying activities across multiple walls. Although one solution is to deploy multiple devices, it is burdensome. To achieve an improved multiple-room classification we demonstrate how the amplitude correlation between *subcarriers #1-5 and #20-25* vary for different human activities. Our empirical results suggest these correlations although fluctuate very slightly due to multipath fading remain similar within the same human activity performed. Hence a significant change ($\rho \geq \pm\ 0.5$) in correlation pattern corresponds to a change in human activity and until the activity is continuously done within the same environment, ρ remains $\leq \pm\ 0.5$. Thus we use this as a feature to identify a transition in human activity. In WiHACS we adopt both phase and amplitude of subcarriers and leave out phase difference for future exploration.

Due to no standard dataset available we performed the algorithms mentioned in [15] on our dataset and used it as a benchmark to evaluate WiHACS. Our preliminary results suggest WiHACS can detect human activities with a similar accuracy to [15] in LOS and one-wall environments but perform relatively better across a 2-room environment. In future work we plan to carry out further correlation analysis in richer environments and measure WiHACS' robustness in unknown environments.

The remainder of this paper is organized as follows: Section II describes the CSI theory. The signal processing and feature extraction techniques is explained in Section III and IV respectively. Section V explains the multi-class classifier and VI discusses its performance. We conclude our paper and propose future directions in section VII.

II. Channel State Information (CSI)

Most COTS MIMO Wi-Fi devices employ IEEE 802.11n/ac that utilizes Orthogonal Frequency Division Multiplexing (OFDM) at the physical layer. The Channel State Information (CSI) estimates the quality of the communication link influenced by multipath propagation. Denoting the complex valued channel frequency response (CFR) as $H(f,t)$, the transmitted and received signals in frequency-domain is expressed as $Y(f,t) = H(f,t)$ x $X(f,t)$ for every time t. This CFR can be extracted from CSI measurements made at the Intel 5300 NIC per subcarrier. The CSI of an OFDM subcarrier is given as -

$$h = |h|e^{j\psi} \tag{1}$$

where $|h|$ and ψ represent amplitude and phase respectively. If N_t, N_r denote number of transmit and receive antennas, each measurement of complex valued CSI is composed of 30 matrices with dimensions N_t x N_r. The matrix entries correspond to CSI values between an antenna pair at a particular subcarrier frequency at each time. Hence CSI time-series consist of 30 x N_t x N_r streams.

A. Phase Extraction & Calibration

Before a packet is transmitted, due to lack of timing and phase synchronization between transmitter and receiver, the phase ψ is buried in noise and can be expressed as:

$$\tilde{\psi}_m = \psi_m - 2\pi\frac{j_m}{N}\zeta + \alpha + Z \tag{2}$$

where $\tilde{\psi}_m$ is the raw noisy phase for $m = 1:30$ subcarriers indexed according to $j_m = [-28, -26, ...28]$ in a 20 MHz channel width of IEEE 802.11n; N = FFT size (64), ζ is timing offset at the NIC, α is an unknown phase offset and Z is a measurement noise. To calibrate the phase we perform a linear transformation as suggested by [19] to estimate the de-noised phase as -

$$\hat{\psi}_m = \tilde{\psi}_m - \frac{\tilde{\psi}_{30} - \tilde{\psi}_1}{56} - \frac{1}{30}\sum_{k=1}^{30}\tilde{\psi}_j \tag{3}$$

Where $\hat{\psi}_m$ is the calibrated phase of the mth subcarrier. The pseudocode for estimating phase is given in Algorithm I.

Algorithm 1 Pseudocode to extract CSI data

```
Input: HT Packets received at 5300 NIC
Result: Dataset S
for i=1:n packets received do
    get csi_matrix of dimensions N_T x N_R x 30
    for sub=1:30 do
        T_m R_n(sub) ← (m, n) entry of csi_matrix
        Output: amplitude γ & phase ψ
    end for
    Calibrate phase by using (3) for each i-th sample
end for
Output: S = [x_γ1, ..., x_γ30, x_ψ1, ..., x_ψ30], x ∈{T-R pair}
```

III. CSI time series pre-processing

In order to obtain features relevant to human activity classification, we perform the following:

A. 1-D Linear Interpolation

Due to sampling jitter inter-perfect packet delivery is not possible. Hence, we adopt 1-D linear interpolation to align the CSI samples with an even spacing of 1ms.

B. Hampel Filtering

The discrete CSI values for subcarriers contain some abrupt values not instigated by human actions. We utilize the Hampel identifier [17] to remove these outliers by filtering out samples outside the interval $[\mu - \gamma * \sigma, \mu + \gamma * \sigma]$ where μ, γ, and σ represents the mean, standard deviation and mean absolute deviation (MAD) respectively. In our experiments we set window size = 20 and standard deviation threshold = 2. Fig. 1 shows the amplitude of 30 subcarriers before and after Hampel filtering.

C. Subcarrier De-noising

To remove high frequency noises (ex- rates/power adaptation at AP) recent systems applied Butterworth low pass filters [13] or filter out target frequencies by band-pass filters [14]. The limitation is it blunts any rising/falling changes in the data which can be used to detect the start and end of a human activity. To preserve the transitional changes, we adopt 1-dimensional Discrete Wavelet Transform (DWT) based de-noising, which doesn't assume continuities in signals and computationally efficient due to linear time complexity [18].

A wavelet ψ is zero-mean function expressed as -

$$\psi_{a,s}(t) = \frac{1}{\sqrt{s}}\psi\Big(\frac{t-a}{s}\Big) \tag{4}$$

where a and s are the scale and translation parameters respectively. Hence the DWT of series f can be written as -

$$Wf(a,s) = \sum_{k=-\infty}^{\infty} f(k)\frac{1}{\sqrt{s}}\psi^*\Big(\frac{k-a}{s}\Big) \tag{5}$$

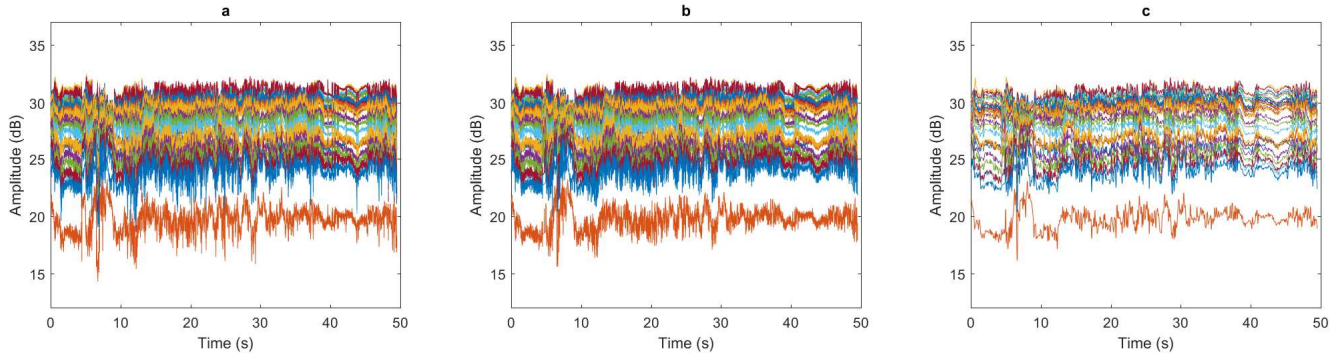

Fig. 1: Pre-processing after phase calibration: (a) Calibrated Phase, (b) Outlier Removal by Hampel Identifier (c) Wavelet De-noising

The DWT of f is estimated by passing samples through banks of low and high-pass filters to produce approximation and detail coefficients respectively. We perform a level-3 Daubechies-4 ('db4') DWT on each subcarrier and use the approximations to reconstruct the signals. From Fig. 1(c) the subcarriers are observed to be cleaner while preserving transitional changes.

D. De-trending OFDM subcarriers

To avoid long-term trends in amplitudes/phases of subcarriers, de-trending was done to avoid spectral energy leakage into lower frequency bands. This is crucial when computing FFT features as slow moving activities (sit, lie down, etc.) occupy very low frequency bands [14].

IV. Data Collection & Feature Extraction

In most related work only one transmit and two receive antennas are utilized and channel is measured only across one transmit-receive pair. Our experiments suggest it is more efficient to measure the subcarriers in all T-R links. To compromise between classification performance and computation we utilize four T-R links from a 2x2 MIMO.

A. Statistical Features

For both amplitude and phase data, a window of length 20 is applied and slid every 5 data points. For every window the *mean, maximum, minimum, median, variance, standard deviation, range, interquartile range, skewness* and *kurtosis* is calculated.

B. FFT Features

1) Normalized Entropy: The *normalized entropy* estimates the disorder measure of time-series data in frequency domain. Denoting N=20, R_i as FFT coefficients (normalized), then normalized entropy is measured as -

$$H = -\sum_{j=1}^{N/2} R_i log_2(R_i) \tag{6}$$

2) Peak Frequency: The peak frequency corresponds to the frequency associated with the highest power when computing power spectrum density (PSD) within a window. The PSD for a signal x of window length L is expressed as-

$$PSD(f) = \frac{1}{L}\sum_{j=0}^{L-1}\Big[(x_j cos(\frac{2\pi f_j}{L})\Big]^2 + \Big[x_j sin(\frac{2\pi f_j}{L})\Big]^2 \tag{7}$$

Since slower $(stand\ \ from\ \ sitting,\ walking)$ and faster $(squatting)$ moving activities occupy $[0, 5Hz]$ and $\geq 5Hz$ as reported in [14], these FFT features aid to distinguish between various daily activities.

C. Correlation Features

For the same window length the correlation between labelled subcarriers is recorded for each T-R link. We observe empirically correlation patterns for different human activities performed in the same environment differ. This will be a key feature as it improves human activity classification specially across multiple rooms as it aids to detect start and end of an activity.

V. Multi-Class SVM

Given training samples (x_i, y_i), where $x_i \in R^n$, i = 1, 2, ...,m, and $y \in (-1, 1)^m$ for a binary classification, a q-class classification can be solved using $q(q-1)/2$ binary classifiers ("one-against-one"). In SVM, the following optimization problem is solved-

$$\underset{w,b,\epsilon}{min} \quad \frac{1}{2}w^T w + C\sum_{i=1}^{m}\epsilon_i \tag{8}$$

$$subject\ to \quad y_i(w^T\phi(x_i) + b) \geq 1 - \epsilon_i, \tag{9}$$

where C is cost parameter $(C, \epsilon_i \geq 0)$ and $\phi(x_i)$ maps x_i onto a higher dimensional space. We used the radial basis function (RBF) as kernels given below -

$$K(x_i, x_j) \equiv \phi(x_i)^T\phi(x_j) \ = exp(-\gamma||x_i - x_j||^2) \tag{10}$$

where γ adjusts the decision boundary. By applying 10-fold cross validation and grid search, optimal values for parameters C and γ were computed. We trained and tested our multi-class SVM by utilizing the LibSVM package [20].

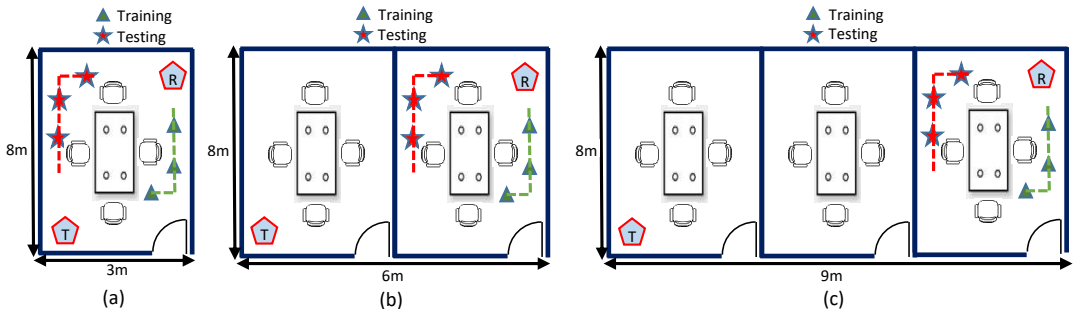

Fig. 2: Three experimental settings used for training and testing purposes: (a) One-room scenario, (b) Two-rooms to emulate a NLOS, (c) Three-rooms to emulate multiple walls.

VI. Experiments & Evaluations

A. Hardware Setup

The transmitter was an Acer 802.11n Dual-band Access Point and the receiver a Dell Latitude with Intel 5300 NIC. We set packet rate to 100Hz, operating frequency at 5GHz, channel width to 20MHz and measured the CSI across 2-transmit & 2-receive antennas.

B. Environmental Settings

The data was collected by volunteers (aged 18-25) in identical side-by-side project rooms. The activities were performed in positions and areas labelled by markers and 'dashed' lines respectively, shown in Fig. 2. A total of 36 batches of data (2 batches per marker location) were collected and each batch consisted of $\approx$ 50 secs of human activities as listed in Table I. The length of each activity was $\approx$ 8s except for fall events. Our fall events only consisted of a volunteer falling sideways on a floor mattress. The sequence of activities performed were random except for $fall \rightarrow lie\text{-}down \rightarrow stand\ \ from\ \ lie\text{-}down$ sequence.

Wi-HACS ACCURACY COMPARISON (%)						
ACTIVITY	Setting #1		Setting #2		Setting #3	
	Wi-HACS	Benchmark [15]	Wi-HACS	Benchmark [15]	Wi-HACS	Benchmark [15]
SIT	**96%**	96%	**92%**	92%	**76%**	69%
STAND	**98%**	100%	**91%**	91%	**74%**	66%
WALK	**98%**	96%	**96%**	93%	**77%**	72%
SQUAT	**97%**	97%	**95%**	92%	**79%**	68%
FALL	**95%**	92%	**90%**	86%	**70%**	59%
LIE-DOWN	**96%**	94%	**88%**	88%	**77%**	68%
STAND FROM LIE DOWN	**96%**	94%	**91%**	86%	**69%**	60%
AVERAGE	**96.6%**	95.5%	**91.8%**	89.7%	**74.5%**	66%
One-tail t test P(T<=t)	**0.18**		**0.09**		**0.001**	

TABLE I: Illustrating WiHACS' performance

C. Performance

We repeated the techniques explained in [15] to compare WiHACS' performance. From Table I it is observed WiHACS perform relatively similar to our benchmark in LOS and one-wall with an average accuracy of 96.6% and 91.8% respectively. However, with a richer environment as setting #3, WiHACS performs with an increased accuracy of 9% compared to the benchmark design. Although for setting 1 there is no much improvement, for setting #2 & 3, the performance improvement is statistically significant at the 90% and 99% confidence level respectively. From our analysis we observed the rate of false positives depend on the sequence of human activities performed. For instance, features collected for 'stand from lying down' position is very similar to a 'stand from sitting' activity. This is primarily why we performed different sequences of activities to ensure a better testing accuracy of our classifier.

Although our correlation analysis is valid for settings #1 & 2, it shows some limitation in setting #3. We observe these correlation patterns do not necessarily stay $\leq \pm 0.5$. However, we do observe a bigger change in ρ when a slow activity is preceded by a faster activity or vice-versa. Therefore the correlation analysis was still utilized as it enabled to identify a change in human activity. In our future work we will include more human activities and study how the correlations behave among the subcarriers.

VII. Conclusion

Human activity classification by wireless systems can facilitate many practical applications. A challenge remains when the classification is done across multiple walls as the signal propagations cannot be effectively predicted in advance. Our proposed system, WiHACS can measure human activities across multiple walls with an improved performance. However, we haven't fully capitalized the parameters in a wireless channel that can be controlled such as transmit power, rates, packet sampling rates etc. In future work we aim to vary these and investigate correlation between subcarriers in richer complex environments. We also aim to include more human activities and measure WiHACS' robustness in identifying finer-grained activities.

Acknowledgment

We like to thank our volunteers who helped collect our dataset. This research is partially supported by the National Research Foundation, Prime Minister's Office, Singapore under its IDM Futures Funding Initiative.

REFERENCES

[1] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, "PADS: Passive detection of moving targets with dynamic speed using PHY layer information," *2014 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, 2014.

[2] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical Human Activity Recognition Using Wearable Sensors," *Sensors*, vol. 15, no. 12, pp. 3131431338, Nov. 2015

[3] G. Kale and V. Patil, "A Study of Vision based Human Motion Recognition and Analysis", *International Journal of Ambient Computing and Intelligence*, vol. 7, no. 2, pp. 75-92, 2016.

[4] O. Waltari and J. Kangasharju, "Content-Centric Networking in the Internet of Things," *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 2016.

[5] N. Patwari and J. Wilson, "Spatial Models for Human Motion-Induced Signal Strength Variance on Static Links," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 791802, 2011.

[6] J. Wilson and N. Patwari, "Radio Tomographic Imaging with Wireless Networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 5, pp. 621632, 2010.

[7] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl, "RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals", *IEEE Transactions on Mobile Computing*, 13(4):907920, 2014.

[8] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," *Proceedings of the 19th annual international conference on Mobile computing & networking - MobiCom '13*, 2013.

[9] B. Kellogg, V. Talla, and S. Gollakota, "Bringing gesture recognition to all devices", In *Proceedings of Usenix NSDI*, 2014

[10] W. Xi, D. Huang, K. Zhao, Y. Yan, Y. Cai, R. Ma, and D. Chen, "Device-free Human Activity Recognition using CSI," *Proceedings of the 1st Workshop on Context Sensing and Activity Recognition - CSAR 15*, 2015.

[11] D. Zhu, N. Pang, G. Li, and S. Liu, "WiseFi: Activity Localization and Recognition on Commodity Off-the-Shelf WiFi Devices," *2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 2016.

[12] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015.

[13] J. Zhang, B. Wei, W. Hu, and S. Kenhere. "Wifi-id: Human identification using wifi signal", *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 7582, 2016.

[14] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A Real-Time and Contactless Fall Detection System with Commodity WiFi Devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511526, Jan. 2017

[15] Y. Wang, X. Jiang, R. Cao, and X. Wang, "Robust Indoor Human Activity Recognition Using Wireless Signals," *Sensors*, vol. 15, no. 7, pp. 1719517208, 2015.

[16] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, p. 53, 2011.

[17] L. Davies and U. Gather, "The identification of multiple outliers," *J. Am. Statist. Assoc.*, vol.88, no. 423, pp. 782-792, 1993.

[18] S. C. C. G. Mallat and Peyre Gabriel, *A wavelet tour of signal processing: the sparse way*, Amsterdam etc.: Elsevier, 2009.

[19] C. Wu, Z. Yang, Z. Zhou, K. Qian, Y. Liu, and M. Liu, "PhaseU: Real-time LOS identification with WiFi," *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015.

[20] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 27, pp. 127, 2011.