ATLANTIS PRESS
Advances in Computer Science Research, volume 44
3rd International Conference on Wireless Communication and Sensor Network (WCSN 2016)

Contrastive Analysis for Human Activity Recognition Algorithms Using WiFi Signals

Jian Zhou
Nanjing University of Posts and Telecommunications,
Nanjing, China
E-mail: 1514991931@qq.com

Jian Zhou, Han Su, Kai Yu
Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Science,
Nanjing, China
E-mail: Han.su @ Jushri.com, Kai.yu @ Jushri.com

***Abstract*-Human Activity monitoring has become increasingly important and has the potential to support a wide area of applications including elder care, well-being management, fitness tracking and building surveillance. Traditional approaches involve wearable sensors and specialized hardware installations. Compared with these solutions, channel state information (CSI) has its advantage. The algorithm for human detection based on CSI Info has become increasingly important. Some prior WiFi signal based human activity recognition systems have been proposed such as Wisee[8], WiFall[11], Witrack[9], CARM[10]. Different from prior works, we propose a contrastive analysis for the recognition algorithms under different transmission frequencies and activities. Finally, the experimental performance of DTW (Dynamic Time Wrapping) and EMD (Earth Mover Distance) is adopted. Our works show that EMD has a better performance than DTW in most cases.**

Keywords-channel state information (CSI); dynamic time warping (DTW); earth mover distance (EMD); activity recognition

I. INTRODUCTION

Human activity recognition is the core technology that enables a wide variety of applications such as health care, smart homes, fitness tracking, and building surveillance. Traditional approaches use cameras [2], radars, or wearable sensors [1]. However, camera based approaches have the fundamental limitations of requiring line of sight with enough lighting and breaching human privacy potentially. Wearable sensors based approaches are inconvenient sometimes.

In 2009, IEEE 802.11n standard indicates that channel response CSI (State Information) can be obtained through the channel state information (Channel). The parameters are extracted from the receiver [6]. CSI shows each sub carrier phase and amplitude information in the channel. By comparing with RSSI, Different from RSSI as the MAC layer superimposition of multipath signals with fast changing phases, the PHY layer power feature, channel response, is able to discriminate multipath characteristics. In a conceptual sense, channel response is to RSSI what a rainbow (color spectrum) is to a sunbeam, where components of different wavelengths are separated. CSI contains more fine-grained information that can be more comprehensive reflection of the multipath effect in the environment, and thus it has a good potential for application in human activity recognition.

In this paper, we firstly study wireless radio propagation model about Device-free activity identification Using WiFi Signatures. Based on the model, we propose a contrastive analysis between EMD(Earth Mover's Distance) and DTW(Dynamic Time Warping) under different transmission frequencies and activities to distinguish human activity.The movements of human body introduce relative unique multi-path distortions in WiFi signals and this uniqueness can be exploited to recognize human activity. Due to the high data rates supported by modern WiFi devices, WiFi cards provide enough CSI values within the duration of a human to construct a high resolution CSI-waveform for each different activity. To differentiate between human activities, we need to extract features that can uniquely represent those activities to establish the fingerprint database. When the environment changes, such as a man walking into the room, the impact of these wireless phenomena on the wireless signals change, resulting in unique changes in CSI values. Then we exploited some algorithm to compare the CSI values with the fingerprint database.

II. RELATED WORKS

Existing solutions can be grouped into three categories: (1) Received Signal Strength (RSS) based,(2) CSI based, and (3) Software Defined Radio (SDR) based.

A *RSS Based*

Sigg et al. proposed activity recognition schemes that utilize RSS values of WiFi signals to recognize four activities including crawling, lying down, standing up, and walking [12]. They achieved activity recognition rates of over 80% for these four activities. Owing to RSS values only provide coarse-grained information about the channel variations and do not contain fine-grained information about small scale fading and multi-path effects caused by thesemicro-movements.

B *CSI Based*

E-eyes recognize a set of nine daily human activities using CSI. Note that WiHear and E-eyes use CSI in quite different ways than CARM. E-eyes uses CSI histograms as fingerprints for recognizing human daily activities, such as brushing teeth, taking showers, and washing dishes, which

are relatively location dependent. In comparison, CARM uses CSI values based on our CSI-speed and CSI-activity models.[8]

C *SDR Based*

Lyonnet et al. use micro Doppler signatures to classify gaits of human subjects into multiple categories using specialized Doppler radars [13].

This paper explores a novel point in the design space and demonstrates that human activity recognition is possible using the existing channel state information provided by IEEE 802.11n devices and using relatively few wireless links, such as those to existing in-home WiFi devices. We implemented the system using a TP-Link WiFi router and a computer with Intel 5300 WiFi NIC. WiFi NICs continuously monitor variations in the wireless channel using CSI, which characterizes the frequency response of the wireless channel , Let $X(f, t)$ and $Y(f,t)$ be the frequency domain representations of transmitted and received signals, respectively, with carrier frequency f ,The two signals are related by the expression $Y(f,t)=H(f,t)\times X(f,t)$, where $H(f,t)$ is the complex valued channel frequency response (CFR) for carrier frequency f measured at time t. CSI measurements basically contain these CFR values. Let N_{Tx} and N_{Rx} represent the number of transmitting and receiving antennas, respectively. As CSI is measured on 30 selected OFDM subcarriers for a received 802.11 frame, each CSI measurement contains 30 matrices with dimensions N_{Tx}, N_{Rx}. Each entry in any matrix is a CFR value between an antenna pair at a certain OFDM subcarrier frequency at a particular time. Onwards, we call the time-series of CFR values for a given antenna pair and OFDM subcarrier as CSI stream. Thus, there are $30\times N_{Tx}\times N_{Rx}$ CSI streams in a time-series of CSI values.

The fundamental drawback of RSSI is that it fails to capture the multipath effects. To fully characterize the individual paths,the wireless propagation channel exploited CSI values. Channel state information or channel status information(CSI) is information that estimate the channel properties of a communication link [5]. In wireless communication, the transmitted radio signal is affected by the physical environment, for instance, reflections, diffraction and scattering. CSI describes how a signal propagates in the channel combined the effect of time delay, amplitude attenuation.

$$\boldsymbol{y}=\boldsymbol{H}\boldsymbol{x}+\mathbf{n} \tag{1}$$

and phase shift.In frequency domain, the narrowband flat-fading channel with multiple transmit and receive antennas (MIMO) is modeled as

where y is the received vector, x is the transmitted

$$\hat{\boldsymbol{H}}=\frac{\boldsymbol{y}}{\boldsymbol{x}} \tag{2}$$

vector, n is the noise vector and H is the channel matrix. As noise is often modeled as circular symmetric complex normal with n ~cN (0, S), H in the above formula can be estimated as

$$H(f)=[H(f_1),...,H(f_k)] \tag{3}$$

And it also can be denoted as

$$H(f_k)=|H(f_k)|e^{j\angle H} \tag{4}$$

Where $H(f_k)$ is the CSI at the subcarrier of central frequency f_k ,amplitude $|H(f_k)|$,and phase $\angle H$.

And multipath propagation can also be characterized by Channel Impulse Response(CIR),CFR is the Fourier transform of CIR. This CIR is given by the following equation.

$$h(\tau)=IFFT(H(f)) \tag{5}$$

This wireless propagation channel is modeled as a temporal linear filter, known as CIR, Under the time-invariant assumption, CIR $h(\tau)$ is also can denoted as:

$$h(\tau)=\sum_{i=1}^{N}a_i e^{-j\theta_i}\delta(\tau-\tau_i) \tag{6}$$

Where a_i, θ_i, and τ_i are the amplitude, phase, and time delay of the ith path, respectively. N is the total number of multipath and $\delta(\tau)$ is the Dirac delta function. Each impulse represents a delayed multipath component, multiplied by the corresponding amplitude and phase.

III. SYSTEM MODEL

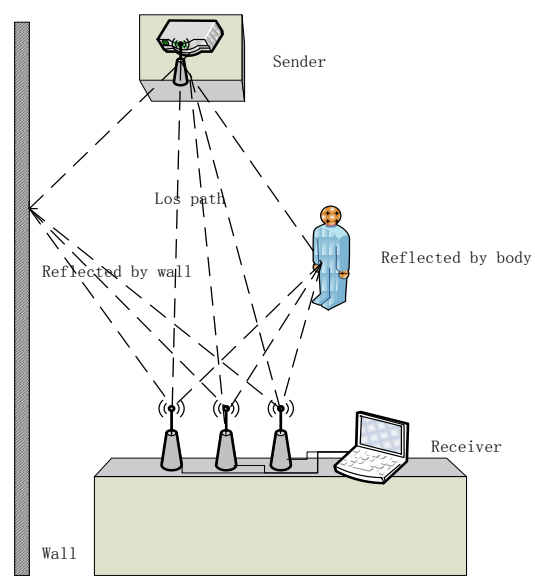


Figure 1. The Experiment scene.

A. System overview

The basic idea of our system is to match CSI patterns against activity profiles. As illustrated In Fig.1, a person do some activities, such as walking, sitting down ,standing up or boxing in the environment of two commercial Off-The-Shelf(COTS) WiFi devices. A sender(such as router)and a receiver(such as a laptop) with 3 antennas as shown in Fig.1.

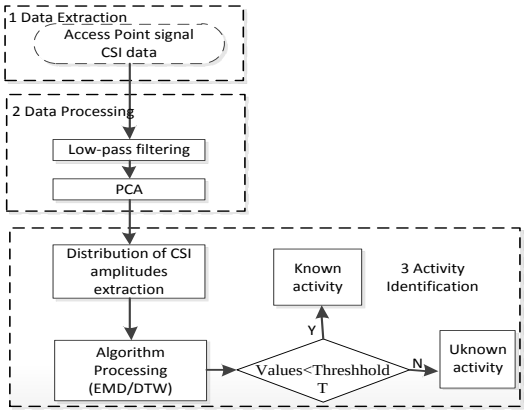


Figure 2. System flow of activity recognition

As shown in Fig2, the system steps are as follows:

1) Data extraction

CSI is collected in thirty subcarriers and three streams, which reflect the signal diversity in frequency and space. We analyze the property of CSI and pick the best ones for detection. What‘s more, CSI is also slightly influenced by environmental noise.

2) Data processing

To use CSI values for recognizing human activities, noise must first be removed from the CSI time series form a low-pass filter to remove high frequency noise, then Principal Component Analysis(PCA) is utilized on the filtered subcarriers to extract the signals that only contain varications caused by activities.

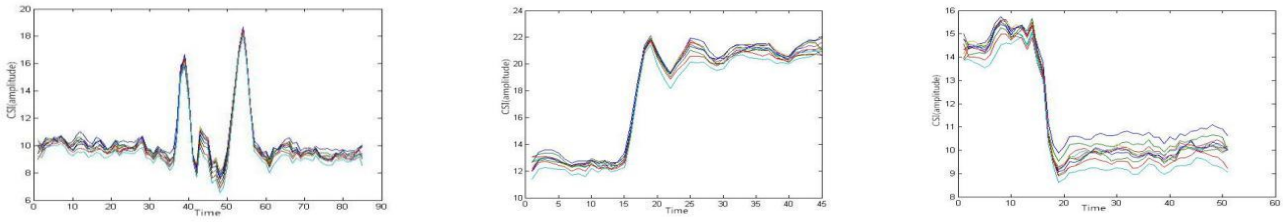
Activity Identification: We can get a set of data that consists of $30\times1\times3$ OFDM subcarriers in a time-series of CSI values .we analyze the OFDM subcarries of CSI and pick the best ones and process it with DTW or EMD for detection. When the distance of two samples is less than the Threshold T, we determine two samples as one action. Otherwise, it is determined as an unknown activity.

Fig.3 shows different kinds of CSI values for three activities, Walking, Standing up and Sitting down. While a man Walking from the receiver to the sender , the waveform of walking consists of many high peaks. The beacon rate is 10 packets per second.

Fig.3 (b) shows a stationary wave before 1.5 second, which the waveform suddenly rises when a person stand up.

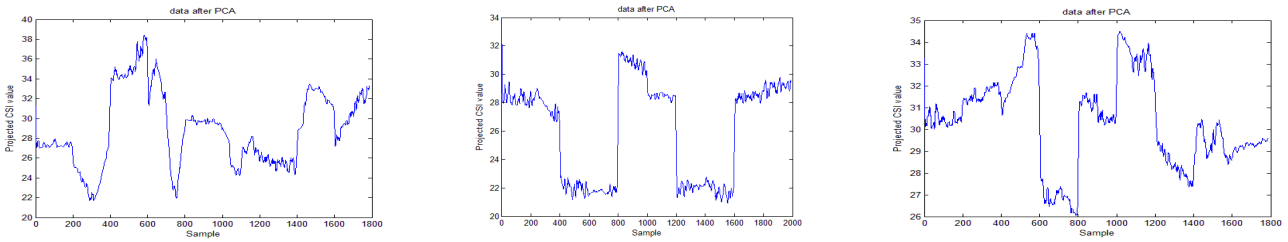
And Fig.3(c) is also play a similarly effects of change. From the experiment result we can see that it is possible to recognize human activities by artificial algorithms.

In order to compare the performance between DTW and EMD, we set the beacon rate as 10 packets per second and 1000 packets per second. As shown in Fig. 4, when we increase the beacon rate to 1000 packets per second, the waveform for waving, boxing, kicking is more fine-grained.



(a)CSI waveform for Walking (b)CSI waveform for standing up (c)CSI waveform for Sitting down

Figure 3. Sitting down, standing up, walking for the beacon rate 10 packets/s.



(a)CSI waveform for waving (b)waveform for boxing (c)CSI waveform for kicking

Figure 4. Waving, boxing, kicking for the beacon rate 1000 packets/s.

B. Dynamic Time Warping

DTW is a dynamic programming based solution for obtaining minimum distance alignment between any two waveform. DTW can handle waveform of different lengths and allows a non-linear mapping of one waveform to another. In contrast to Euclidean distance, DTW gives us intuitive distance between two waveform by determining minimum distance warping path between them even if they are distorted or shifted versions of each other. DTW distance is the Euclidean distance of the optimal warping path between two waveform calculated under boundary conditions and local path constraints[4]. In our experiments, DTW distance proves to be very effective metric for comparing two shape features of different activities, such as walking1 to walking2, walking to running, walking to kicking. In our experiments, we use the open source implementation of DTW in the Machine Learning Toolbox (MLT) by Jang [3].

C. Earth Mover's Distance

The Earth Mover's Distance (EMD) is a method to evaluate dissimilarity between two multi-dimensional distributions in some feature space where a distance measure between single features, which we call the ground distance is given. The EMD "lifts" this distance from individual features to full distributions.[7]Intuitively, given two distributions, one can be seen as a mass of earth properly spread in space, the other as a collection of holes in that same space. Then, the EMD measures the least amount of work needed to fill the holes with earth. Here, a unit of work corresponds to transporting a unit of earth by a unit of ground distance. A distribution can be represented by a set of clusters where each cluster is represented by its mean (or mode), and by the fraction of the distribution that belongs to that cluster. We call such a representation the signature of the distribution. The two signatures can have different sizes, for example, simple distributions have shorter signatures than complex ones[7].We use the Earth Mover's Distance(EMD) to quantify the similarity of two distributions.

IV. IMPLEMENTION AND EVALUATION

The experiments are conducted in two scenarios:(a)different activities(sitting down, standing up, walking)for the beacon rate 10 packets per second, which is 10Hz.(b) different activities(waving, boxing, kicking)for the beacon rate 1000 packets per second, which is 1000Hz. In table1 and talbe2, red indicates error recognition and blue denotes correct recognition. As illustrated in Table 1,the value of Sitting down1, Standing up1, Walking1 is template, and Sitting down2, Standing up2, Walking2 is sample data. The threshold of sitting down is 80.5, when the threshold of DTW values between sitting down1 and sitting down2 is less than 85, the sitting down1 and sitting down2 can be classified one action. Under the same condition, we make the same experiments with EMD.

As show from Table 1, an error which classifies the action of Sitting down as Walking has occurred in DTW. In the same circumstance, the algorithm accuracy of EMD can get 100% and it identifies the error. The histograms of CSI amplitudes (quantized to 6 bins) for subcarrier 25, however, show very distinct distributions that can clearly distinguish these different activities in the same case. The EMD value of different activities has an obvious difference. Compared with DTW algorithm, EMD is more suitable for human activities when set the beacon rate 10 packets/s.

In order to detect the robustness of two algorithms .we do a series of complicated activities and increase the beacon rate to 1000 packets per second, which is 1000Hz. Many errors occurs in DTW testing and two errors take place in EMD testing. It can be proved that the performance of the EMD is better than DTW.

TABLE I. SITTING DOWN, STANDING UP, WALKING FOR THE BEACON RATE 10 PACKETS/S

DTW VALUES	Sitting down1	Sitting down2	Standing up1	Standing up2	Walking1	Walking1
Sitting down1	0	70.257	86.522	129.365	130.725	120.115
Sitting down2	70.257	0	50.030	221.303	48.926	28.235
Standing up1	86.522	50.030	0	31.26	153.362	188.182
Standing up2	129.365	221.303	31.26	0	498.161	521.380
Walking1	130.725	48.926	153.362	498.161	0	7.006
Walking2	120.115	28.235	188.182	521.380	7.006	0
EMD VALUES	Sitting down1	Sitting down2	Standing up1	Standing up2	Walking1	Walking1
Sitting down1	0	223.922	1362.333	401.162	466.465	401.162
Sitting down2	223.922	0	1358.058	555.800	615.187	558.800
Stand up1	1362.335	1358.058	0	190.769	1786.671	1721.581
Stand up2	1426.033	1401.399	190.769	0	1878.882	1813.181
Walking1	466.465	615.187	1786.671	1878.882	0	77.321
Walking2	401.162	558.800	1721.581	1813.181	77.321	0

TABLE II. WAVING, BOXING, KICKING FOR THE BEACON RATE 1000 PACKETS/S

DTW VALUES	Waving1	Waving2	Boxing1	Boxing2	Kicking1	Kicking2
Waving1	0	31.216	198.339	15.73	466.698	1126.147
Waving2	31.216	0	52.927	45.823	230.517	721.469
Boxing1	198.339	52.927	0	83.777	152.851	452.120
Boxing2	15.73	45.823	83.777	0	421.206	1051.972
Kicking1	466.698	230.517	52.851	421.206	0	86.262
Kicking2	1126.147	721.469	452.120	1051.972	86.262	0
EMD VALUES	Waving1	Waving2	Boxing1	Boxing2	Kicking1	Kicking2
Waving1	0	1828.213	2543.691	2911.373	7800.924	7599.317
Waving2	1828.213	0	744.046	2765.910	5981.657	5771.603
Boxing1	2543.691	744.046	0	2470.666	5265.844	5055.889
Boxing2	2911.373	2765.910	2470.666	0	7729.751	7525.685
Kicking1	7800.924	5981.657	5265.844	7729.751	0	790.381
Kicking2	7599.317	5771.603	5055.889	7525.685	790.381	0

V. LIMITATION AND FUTURE WORK

In this paper, we exploit COTS WiFi devices for recognizing human activities to analyze the algorithm performance. However, our current experiment has three limitations. Firstly, this work has an insufficient number of experiments that based on our random experiment. Secondly, the accuracy of our current scheme is affected by human's shape including height and weight. Finally, in the data processing, it is difficult to separate the activities signal from interfering ones absolutely. We will explore the

method of DWT (Discrete Wavelet Transform) or PLCR (Pass Length Change Rate) in the next step.

VI. THE CONCLUSION

This paper demonstrates that human activities recognition can be estimated using off-the-shelf WiFi devices already everywhere in our daily life. We take advantages of the wireless physical information-Channel State Information (CSI) in widely deployed commercial wireless infrastructures. To demonstrate the feasibility and effectiveness, we implemented this model on Linux platform with commercial 802.11n NICs. The experiments were conducted in two different scenarios with DTW and EMD. According to our observations from experiments, in the simple scenario, the DTW and EMD get acceptable detection accuracy, but EMD has a better performance. Although EMD have qualified performance, it has a low differentiating degree. In the next step, we will utilize the algorithms of SVM (Support Vector Machine) and HMM (Hidden Markov Model).

ACKNOWLEDGEMENT

This work is supported by the National Hi-tech Research and Development Program of China (863 Program) under Grant No 2015AA01A705.

REFERENCES

[1] K. Yatani and K. N. Truong. Bodyscope: a wearable acousticsensor for activity recognition. In Proceedings of ACMUbiComp, pages 341–350, 2012.

[2] J. K. Aggarwal and S. R. Michael. Human activity analysis:A review. ACM Computing Surveys, 43(3), 2011.

[3] Jyh-Shing Roger Jang. Machine learning toolbox,available at http://mirlab.org/jang/matlab/toolbox/machinelearning,accessed on december 23, 2014.

[4] Meinard Müller. Dynamic time warping. Information retrieval for music and motion, pages 69-84, 2007.

[5] K. Wu, J. Xiao, Y. Yi, M. Gao, and L. M. Ni, “Fila: Fine-grained Indoor localization,” in INFOCOM, 2012 Proceedings IEEE. IEEE, pages 2210-2218,2012.

[6] Halperin D, Hu W, Sheth A, et al. Predictable 802. 11 Packet Delivery from Wireless Channel Measurements[J]. ACM Sigcomm Computer Communication Rview, pages 159-170, 2011.

[7] http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/RUBNER/emd.htm

[8] Yan Wang, Jian Liu, E-eyes: Device-free Location-oriented Activity Identification Using Fine-grained WiFi Signatures, Mobicom pages 617-628, 2014.

[9] Kamran Ali Alex X. Liu, Keystroke Recognition Using WiFi Signals, Mobicom pages 90-102,2015.

[10] Wei Wang, Alex X. Liunderstanding and Modeling of WiFi Signal Based Human Activity Recognition, Mobicom pages 65-76,2015.

[11] C. Han, K. Wu, Y. Wang, and L. M. Ni. Wifall: Device-free fall detection by wireless networks. In Proceedings of IEEE INFOCOM, pages 271-279, 2014

[12] Stephan Sigg, Markus Scholz, Shuyu Shi, Yusheng Ji,and Michael Beigl. Rf-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals. Mobile Computing, IEEE Transactions on, 13(4):pages907–920, 2014.

[13] Bastien Lyonnet, Cornel Ioana, and Moeness G Amin.Human gait classification using microdopplertime-frequency signal representations. In Radar Conference, 2010 IEEE, pages915–919. IEEE, 2010.