WiDMove: Sensing Movement Direction using IEEE 802.11n Interfaces

Bruno Soares da Silva
Universidade Federal de Goiás
Instituto de Informática
Goiânia - Brasil
brunodasilva@inf.ufg.br

Gustavo Teodoro Laureano
Universidade Federal de Goiás
Instituto de Informática
Goiânia - Brasil
gustavo@inf.ufg.br

Abdallah S. Abdallah
School of Engineering
Penn State Behrend
Erie, PA – USA
aua639@psu.edu

Kleber Vieira Cardoso
Universidade Federal de Goiás
Instituto de Informática
Goiânia - Brasil
kleber@inf.ufg.br

***Abstract*—The accurate detection of people in indoor environments requires high-cost devices, while low-cost devices, in addition to low accuracy, offer little information about the monitored events. The perturbations that result from indoor movements affect the signals received by 802.11 interfaces. Hence, an 802.11 device becomes a widely available, low-cost, and reasonably accurate solution for several applications. This paper presents WiDMove, a proposed technique to detect the entry and exit of persons, within an indoor environment, using the channel state information (CSI) measurements, which is provided by the IEEE 802.11n compliant devices. Based on the gathered CSI measurements, we utilized frequency-time analysis methodology to build an efficient features vector based on Short-Time Fourier Transform (STFT) and Principal Component Analysis (PCA). We used the extracted features to train and develop a Support Vector Machine (SVM) classifier, which provided very promising initial results. Our initial results have an accuracy near 80%.**

I. Introduction and Related Work

The *IEEE 802.11n* standard has defined a mechanism, known as Channel State Information (CSI), which can monitor the channel of 802.11 interfaces. This mechanism can estimate the changes in the signal when it travels from the transmitter to the receiver. With this mechanism, we can measure the perturbation caused by human activities using off-the-shelf 802.11 interfaces, which are widely deployed.

There are prior proposals to use CSI to detect human activities such as the work presented in [1] and [2]. Authors in [1] utilized CSI measurements to recognize human subjects or various classes of activities using a Naive Bayes classifier. In [2], authors presented quantitative models to be used as matching profiles to correlate CSI measurement with a specific set of activities (e.g., running, walking, falling,and setting-down). The authors used wavelets analysis to extract movement-related features and Hidden Markov Model (HMM) for the classification stage.

The contributions of this paper can be summarized as follows. We propose WiDMove, a new system prototype to use off-the-shelf 802.11n interfaces as sensors to detect specifically the human movement direction based on CSI measurements. WiDMove uses, among other techniques, Short-Time Fourier Transform (STFT) to extract movement corresponding features and support vector machine (SVM) classifier for the classification stage. In addition, we developed and tested the prototype in lab environment using real samples dataset, which we created by collecting samples from eight different persons. The dataset used in developing this system will be made available for reuse by research community upon request.

II. Background

Channel state information (CSI) is a channel monitoring mechanism implemented in the IEEE 802.11n compliant interfaces, which can describe the changes in the amplitude and phase occurred in a transmitted signal during the transmission process [3]. This mechanism was proposed with the desire to support the 802.11 devices in the environment adaption process. However, the information offered by this mechanism started to be used for various proposes since it correlates with reasonably promising accuracy with the physical activities happening at the surrounding environment.

When a room remains without changes, i.e, don't have movement within, the changes to the CSI measurements are negligible. However, when some movement activities occur in same room, the changes to the CSI measurement are observable and quantifiable, due to the multipath effect on the received signals.

IEEE 802.11n devices use OFDM modulation [4], therefore the CSI metrics include the description of changes in each channel sub-carrier. A block of CSI measurement has the dimensions of $N_{Tx} \times N_{Rx} \times N_{Sub}$, where N_{Tx} and N_{Rx} are the number of transmitting and receiving antennas in order, and N_{Sub} is the number of sub-carriers reported by the device. The CSI measurements are updated with every new packet arrives at the receiver according to certain conditions such as being transmitted in High-Throughput (HT) mode [3].

In this paper, we use support vector machine (SVM) classifier. In classification problems, the classifier's model must be trained using pre-labeled dataset. To create this model, it is necessary to represent the data into usually a compressed format compared to the original acquired measurements to reduce the system complexity. This is done by extracting a set of of features, that can describe the most important attributes of the desired classes, while maximizing separability metrics between classes. The main signal processing and pattern recognition techniques, we used in this paper, include principle component analysis (PCA), Short-Time Fourier Transform (STFT), and Support Vector Machines (SVMs). For more

978-1-5386-2410-4/18/$31.00 ©2018 IEEE

details on their technical details, the readers may review [5], [6], and [7] respectively.

III. WiDMove: System Overview

A. Environment Setup

An electromagnetic signal can be perturbed much more by objects whose dimensions are larger than the signal wavelength. The wavelength is defined by $\lambda = c/f$, where c is the light speed and f is the channel frequency in Hertz. In this setup, we use 5 GHz channels, therefore the signal can be perturbed much more by objects that has dimensions larger than 6 cm. Because of this sensitivity, it is important to take some measures to minimize the impact of environment noise. We take advantage of the use of directional antennas by positioning a pair of directional antennas at the entry door's extremities. That way, the signal should not get significantly perturbed unless a person passes through the door resulting into obstructing the line of sight (LoS) [8].

We built an experimental setup in lab using two computers equipped with Intel Core i7 processors, 8 GB of RAM, and Ubuntu 14.04. Both computers are equipped with Atheros ath9k AR9380 NICs, which implements the 802.11n standard and therefore has support to CSI. The tools to extract the CSI measurements was installed in both computers, as specified by the developers [9].

The interfaces are configured to run an infra-structure network between the computers using one channel on the baseband of 5 GHz with a bandwidth of 20 MHz.

B. Data Collection

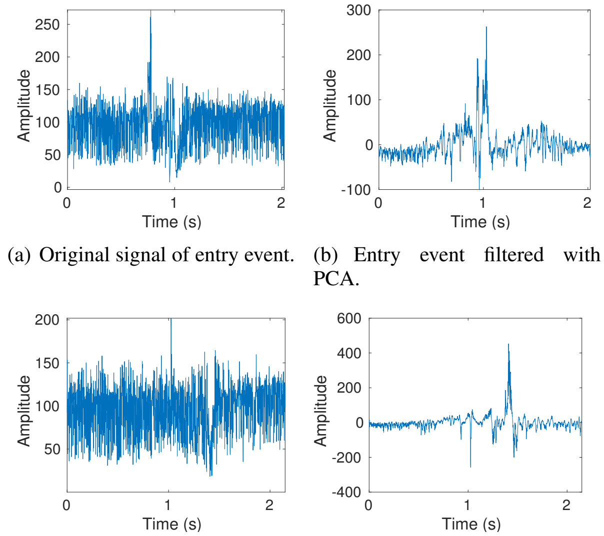
We performed the tests with 8 persons, 7 males and one female, with ages between 19 and 27 years. We collected 264 samples for each event type (i.e., entering and existing) where the peoples were advised to walk normally while the monitoring setup is running. The data were then formatted, categorized, and labeled to build the WiDMove dataset, which is made available for download by the research community upon request.

As the human body activities can be described at a frequency of up to 300 Hz, we sample the CSI measurements at a sample rate of 800 pps. We organize the CSI data as a matrix with dimension $(800 \times Secs) \times N_{Tx} \times N_{Rx} \times N_{Sub}$, which describes $N_{Tx} \times N_{Rx} \times N_{Sub}$ CSI streams in time domain. We only use the amplitude portion of the CSI measurements because of the lack of accuracy and consistency in the Carrier Frequency Offset (CFO), which makes the phase data reported by the available API not reliable [2]

C. Data Processing

We start the filtering process by rearranging the CSI matrix on a bidimensional matrix that describes the amplitude of CSI streams over the time. Then, we remove the LoS energy from each point of CSI streams by subtracting 4 seconds mean from each CSI stream. This is followed by applying the CSI denoising technique based on PCA [2]. The technique uses the correlation of each CSI stream to pick the components that better describe the human activities. Figure 1 shows four signals, that represent both entry and exit events before and after the denoising filter process.

At the end of this step we have a matrix with dimensions $(N_{Tx} * N_{Rx} * N_{Sub}) \times (800 * N_{Secs})$, which represents the PCA components amplitude ordered by the highest representativeness over the time.



(a) Original signal of entry event. (b) Entry event filtered with PCA.

(c) Original signal of exit event. (d) Exit event filtered with PCA.

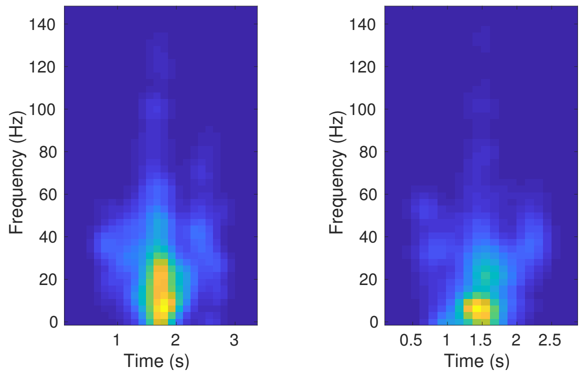
Fig. 1: Comparison of signals before and after PCA filtering.

D. Spectrogram creation

In this section, we describe the creation of the spectrogram that we use to extract the features required for classifying entry and exiting events. Even after the PCA denoising, the components still have informations that don't correspond to the desired human activity, therefore, we need some techniques to enhance the creation of this type of spectrogram, such as the technique proposed in [10], which we use as basis for our solution.

We start by computing the STFT from the first 20 PCA components using a window size of 256 samples and 155 overlapped samples. These configurations provide us with a resolution level of 3.12 Hz and 0.12 seconds in the STFT window.

For each STFT chunk obtained, we execute 3 steps for preprocessing: (1) Zero the energy of frequencies above 146 Hz. (2) Zero the energy of each silenced time chunk that doesn't reach the minimum energy threshold, which we obtained through calibration measurements. (3) Normalize the energy from each non-silenced chunk by dividing by the chunk total energy and the subtracting the average frequency energies. After preprocessing, we get the final spectrogram by overlapping all the preprocessed STFTs chunks, then summing each corresponding point. The output of the summing step is then smoothed by using a 2-D Gaussian filter with $\alpha = 0.8$ and $size = 5$. Figure 2 shows the final spectrogram examples for both an entering and an exiting event.



(a) Spectrogram of an entering event.

(b) Spectrogram of an exiting event

Fig. 2: Spectrograms of entry and exit events.

E. Detecting an active event period

This section describes our technique to detect the period of an active event based on the spectrogram, which we built as described earlier in Section III-D. In comparison to previous techniques that use only the variations in time domain to detect events, this technique is more resilient to environment changes because it also takes into account the changes in frequencies, which allows to filter events according to energy levels in frequency domain.

Entry and exit events are composed mainly of movements by the human torso area. That movement normally occurs in an approximate velocity of 1 m/s. Using the formula $f = 2v/\lambda$ described in [10], we can check whether a movement exists or not, in a 5 GHz channel, in a frequency of 33 Hz. Because of that, we use the spectrogram frequency range from 20 to 50 Hz to detect the exit and entry events.

The interval definition process starts by scanning all time intervals in the spectrogram for a minimum duration of 1.2 seconds that have energy in all chunks. Then, we check in each obtained interval if the sum of the interval energy normalized by the division of the average energy of all intervals is superior than a specific threshold, which we obtained through calibration measurements during the data collection stage to build our benchmarking dataset. If the interval is in accordance with these restrictions, we consider it as a single interval of entry or exit event.

F. Features extraction & classification

The first feature extracted is the event duration in seconds. Then, we generate three signals that describe the accumulated energy percentage of the spectrogram in the event interval, using the 25, 50 and 95%. These signals are generated by extracting, from each time chunk, the frequencies where the accumulated energy reaches the desired percentage, as defined by equation 1, where $f_p(p,t)$ represents the frequency in chunk t, at which the percentage p was reached, f_{max} represents the maximum frequency of the spectrogram and $E_{(t,x)}$ represents the energy of frequency x in chunk t. After generating these three signals, we re-sample each one into a signal that has 30 points, using the nearest neighbors interpolation method. This results into a set of 90 features, named as the energy percentage features, as illustrated at the horizontal axis of figure 3.

$$f_p(p,t) = \min_{\forall f \in [1,\dots f_{max}]} \left\{ f \;\middle|\; \frac{\sum_{a=1}^{f} E_{(t,a)}}{\sum_{x=1}^{f_{max}} E_{(t,x)}} \geq p \right\} \quad (1)$$

The human torso and legs movements can be described by the percentile accumulative energy signals from 50 and 95% respectively [11]. For each value present in this signals, we compute the movement speed using the formula $v = f * \lambda/2$ [10]. For each one of this two signals, we compute seven values: mean, variance, minimum, maximum, skewness, kurtosis and entropy. By the end of this step, we get an additional 14 features set, named as the speed features, as shown by figure 3.

The next features set describes the event energy signature, which is computed by dividing the event interval in 4 chunks then calculate, for each chunk, the energy average of each frequency found in the spectrogram (In this case, 33 frequency windows). This results into a new set of 132 features, which we call as the energy signature features in figure 3.

Finally, we extract the features in the second PCA component that can better describe the human activity [2]. We start by filtering the component using a Continuous Wavelet Transform (CWT) applied with Morlet Wavelet and them returning it to the time domain using the ICWT with parameters to get the frequency range from 0 to 100 Hz. We divide that signal in 10 chunks and compute 10 values: mean, variance, minimum, maximum, zero crossing rate, skewness, kurtosis, entropy, energy mean and the frequency peak obtained by the FFT transformation. As result, we get the last 100 features of the event, named as PCA components in figure 3.

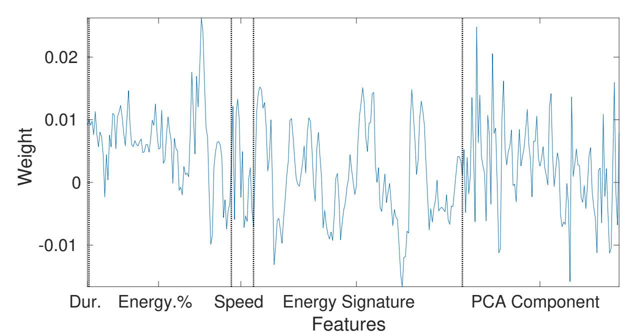
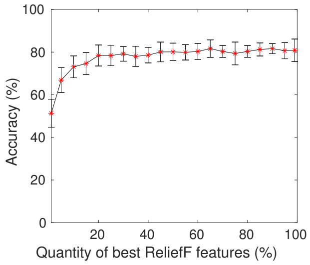
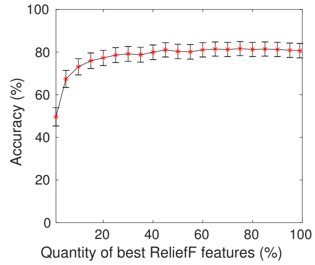


Fig. 3: Subtypes of our features vector (i.e., energy percentiles, speed, energy signature, pca component) and their corresponding weights as determined by the ReliefF algorithm

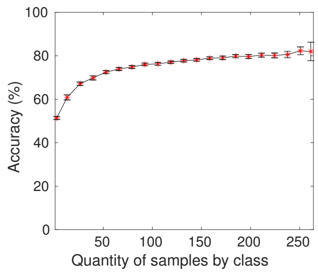
To create the model to identify the movement event direction, we created a Support Vector Machine (SVMs) classifier with a kernel whose function is Radial Basis Function (RBF). We have implemented all signal processing, classifier training, and testing stages in Matlab®.



(a) Average accuracy of classifiers using K-Fold cross-validation (CI of 95%).



(b) Average accuracy of classifiers using Leave-One-Out cross-validation (CI of 95%)



(c) Impact of samples used in training phase (CI of 95%).

Fig. 4: Experiment results.

IV. Performance evaluation

For validating the SVM classifier, we used Leave-One-Out and K-Fold cross-validation techniques, with $K = 10$ for the latter. The results shows that the WiDMove can hit an accuracy near 80%, under a Confidence Interval (CI) of 95%, as shown in figure 4.

A. Impact of features quantity

Since the amount of features and their combination may have an influence on both performance and accuracy, hence we use the ReliefF algorithm to rank the features that we extracted to determine which subset might be more efficient in building the desired classifier. Figure 3 shows the weights of each feature used in the WiDMove. After finding the weights of features, we create classifiers that use higher weight features, varying the quantity of features up to the total set. We evaluate these classifiers using cross-validation techniques (e.g., K-Fold and Leave-One-Out). The results we acquired show that the use of around 40% of features can result in an accuracy near 75%, as shown in figures 4(a) and 4(b).

B. Impact of samples used in training phase

Finally, using a classifier that consumes the entire 337 features vector, we quantified the impact of the size of training set. We created a large set of identical classifiers varying the quantity of samples used in the training phase. We checked the average accuracy for 50 random executions for each classifier. As shown by figure 4(c), the classifier that got trained with 100 samples has an accuracy higher than 70%. This accuracy was increased by adding more samples to the training phase, resulting into an accuracy near 80%. Classifiers were only tested using samples from outside the training subset. Accuracy has been calculated according to equation 2.

$$Accuracy = \frac{TruePositive + TrueNegative}{TotalPopulation} \qquad (2)$$

V. Conclusion

In this paper, we presented WiDMove, a new technique that uses the IEEE 802.11n CSI measurements and support vector machines (SVMs) classifier to detect the walking direction of humans in an indoor environment. The new technique achieved an average accuracy near 80% with a features vector of only 337 features, which is a reasonable solution in terms of computational complexity, real-time requirements, ease of use and wide spread of deployment. In addition, we briefly summarized the description of the WiDMove's CSI measurements dataset for indoor human walking activities, which we collected, formatted, and labeled. The WiDMove dataset is available for research purposes upon request. Currently, we focus on investigating other time-frequency analysis methods and off-the-shelf (OTS) devices to address the challenge of inaccurate and inconsistent CSI phase measurements. Alternative machine learning classifiers are also to be investigated with the goal of increasing the system accuracy.

References

[1] Y. Li and T. Zhu, "Using wi-fi signals to characterize human gait for identification and activity monitoring," in *Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2016 IEEE First International Conference on*. IEEE, 2016, pp. 238–247.

[2] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking*. ACM, 2015, pp. 65–76.

[3] "Ieee standard for information technology– local and metropolitan area networks– specific requirements– part 11: Wireless lan medium access control (mac)and physical layer (phy) specifications amendment 5: Enhancements for higher throughput," *IEEE Std 802.11n-2009*, pp. 1–565, Oct 2009.

[4] "Ieee standard for telecommunications and information exchange between systems - lan/man specific requirements - part 11: Wireless medium access control (mac) and physical layer (phy) specifications: High speed physical layer in the 5 ghz band," *IEEE Std 802.11a-1999*, pp. 1–102, Dec 1999.

[5] I. T. Jolliffe, *Principal component analysis*. Springer, 2002.

[6] Y. Zhang, Z. Guo, W. Wang, S. He, T. Lee, and M. Loew, "A comparison of the wavelet and short-time fourier transforms for doppler spectral analysis," *Medical Engineering & Physics*, vol. 25, no. 7, pp. 547 – 557, 2003.

[7] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, Mar 2002.

[8] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *Distributed Computing in Sensor Systems (DCOSS), 2016 International Conference on*. IEEE, 2016, pp. 75–82.

[9] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wifi," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '15. New York, NY, USA: ACM, 2015, pp. 53–64. [Online]. Available: http://doi.acm.org/10.1145/2789168.2790124

[10] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 363–373.

[11] P. Van Dorp and F. Groen, "Feature-based human motion parameter estimation with radar," *IET Radar, Sonar & Navigation*, vol. 2, no. 2, pp. 135–145, 2008.