Non-intrusive Biometric Identification for Personalized Computing Using Wireless Big Data

Zhiwei Zhao[1], Zifei Zhao[1], Geyong Min[2,*] and Chang Shu[1]
[1]School of Computer Science and Engineering, University of Electronic Science and Technology of China, China
[2]Department of Computer Science, University of Exeter, UK
Email: zzw@uestc.edu.cn, {zifei,shuchang}@mobinets.org, g.min@exeter.ac.uk

***Abstract*—Human identification is of great importance for personalized and ubiquitous computing services. As the rapid and ubiquitous deployment of wireless access, the data of channel state information (CSI) gathered from wireless networks has become a useful way for recognizing human activities. In this paper, we investigate the problem of human identification based on biometrics using the wireless big data of CSI. The rationale is that human's behavioral movements cause unique impacts on wireless CSI, which can be used to recognize the corresponding biometrics and further identify different persons in a *non-intrusive* manner. We first propose BioID, a general frame for biometric identification based on wireless big data of CSI. Based on the framework, we further devise a novel scheme for identification with lip motions using CSI. By considering the CSI characteristics of lip motions, we employ various signal processing and classification techniques. The experimental results show that the proposed scheme can achieve accurate human identification (accuracy >90%). Finally, we discuss several open issues and challenges for future direction.**

***Index Terms*—Channel state information, behavior recognition, biometrics, identification,**

I. INTRODUCTION

Human identification is critical for personalized computing and services. The traditional human identification strategies such as password, personal identification numbers (PIN) and on-screen gestures are prone to security risks since they can be easily copied, shared or peeped. Biometrics, as a method for human identification, have been extensively studied in recent years and have replaced the traditional identification strategies in various fields due to their distinctive nature [1]. For example, fingerprint has been widely applied in the building access, smartphone authentication, residence permits, criminal records, etc. There are plenty of different biometrics, which are typically clustered into two categories: Physiological and Behavioral, as shown in Figure 1. Physiological biometrics include faces, fingerprint, ears, palm veins, thermograms, body odor, DNA, etc. Behavioral biometrics include signature, voice, keystrokes, walking poses, eye motions, lip motions, etc. To exploit the above biometrics for human identification, various devices are designed and deployed such as fingerprint scanner, facial recognition cameras and palm print systems. These devices are often placed in front of the door access and require the users to proactively measure their biometrics for identification. The devices however are often expensive and require considerable deployment overhead.

Recent advances in wireless communications shed light on the design of wireless non-intrusive biometric identification [2]–[4]. Specifically, the channel state information (CSI) is promising to "infer" the biometrics with the channel states variations interfered by people's behavioral movements. Modern Wi-Fi standards like IEEE 802.11n/ac convey massive CSI data since they employ Multiple-Input Multiple-Output (MIMO) and Orthogonal Frequency Division Multiplexing (OFDM) to improve the link capacity of the wireless communication. In such systems, there are multiple subcarriers to transmit the wireless signal at the same time. CSI measures the channel frequency response (CFR) in the subcarriers. The received signal is a result of the constructive and destructive interference of several multi-path signals scattered from the walls, pedestrians and other surrounding objects. Therefore, the subtle physical movement can be easily reflected by the CSI. Recently, CSI has been identified as a key enabling characteristic for the human-machine interaction and has now been measurable on the commercial-off-the-shelf (COTS) WLAN infrastructures such as the Intel 5300 NIC [5] and Atheros series [6].

From the nature of CSI, it can be inferred that the behavioral biometrics are more detectable than the physiological biometrics. The reason is that the behavioral motions can affect the wireless signals with certain unique patterns, which can be possibly captured via CSI and recognized at the receiver side. For example, Li et al. [7] used CSI traces to extract the distinct gestures which could be used for human identification. Xin et al. [8] extracted the walking patterns of individuals from CSI measurement, which can further distinguish different people. The challenges for the exploitation of CSI for biometric identification include: 1) Extraction of the signal reflections from the target motions. The behavioral biometrics are detected by the signal variations. However, the movements of the surrounding objects and people can also affect the signal reflections. As a result, it is very challenging to clean the noise data and extract the signal reflections caused only by the motions of the target behavioral biometrics. 2) Recognition of the tiny differences among the biometrics of different people. Considering that the typical features like gaits, talking rates may be similar for different people, it is challenging to further extract the distinct biometric patterns for different individuals. To address this problem, appropriate features regarding the biometric nature need to be identified and extracted. 3) There may be multiple data sources that can be used at the same time, the aggregation of these different data and features is also challenging for accurate biometric recognition.

In this paper, we discuss the wireless biometric identification using the CSI data. We first introduce the fundamental ideas of the biometrics and CSI-based applications. Then we propose a general framework for wireless biometric identification based on CSI big data, which covers the current literature and is extensive to support recognition for more types of biometrics. With the framework, we further devise a novel scheme for wireless biometric identification of lip motions. We carefully design the noise removal and feature extraction modules in the lip motion based identification and implement it on Intel 5300 NIC. The experiments with volunteers show that the scheme can accurate identify different users using the CSI data. The major contribution of this paper include:

1) We propose a general framework for biometric-based human identification using wireless CSI big data. The framework can be used to derive various identification schemes with different biometrics.
2) Based on the framework, we propose a novel lip-motion based identification approach. The CSI characteristics associated with lip motions are effectively considered in the noise removal and feature extraction.
3) We implement the lip-motion based identification. The experimental results show that the proposed work can accurate identify different people based on the CSI data.

The rest of the paper is organized as follows. Section II introduces the basic concepts of biometrics and the applications of CSI. Section III presents the general framework for biometric identification with wireless big data. Section IV presents the detailed design of the lip motion based identification, which is an instance of the framework. Section V evaluates the proposed work. Section **??** discusses open issues and challenges for future works. Section VI concludes this work.

II. Biometrics and Channel State Information

In this section, we present the preliminaries on biometrics and CSI. The fundamentals of biometrics and CSI are summarized. Then the related works on wireless biometrics identification based on CSI are reviewed and discussed.

A. Biometrics

Biometrics are the distinct characteristics of individuals, such as face, fingerprint, etc. Biometric IDs have been widely used in building access, system authentication, national ID, residence permits, personalized services, etc [1].

As shown in Figure 1, biometrics can be basically classified into two categories: Physiological and behavioral. Typical physiological biometrics include the faces, fingerprints, palm veins, etc. Typical behavioral biometrics include keystrokes, gaits, gestures, lip motions [9], [10], etc. The current literature focuses on the use of physiological biometrics for identification, which require various types of specialized devices such as fingerprint scanners, face detection cameras, etc., which are often expensive and require extra deployment overhead. There are some new works aiming at exploiting the behavioral biometrics using the cameras [11]. One limitation of the camera-based works is that the target users should be in the line-of-sight and with good light condition. Otherwise, the graphs cannot be captured in good quality and the biometric identification cannot be done.

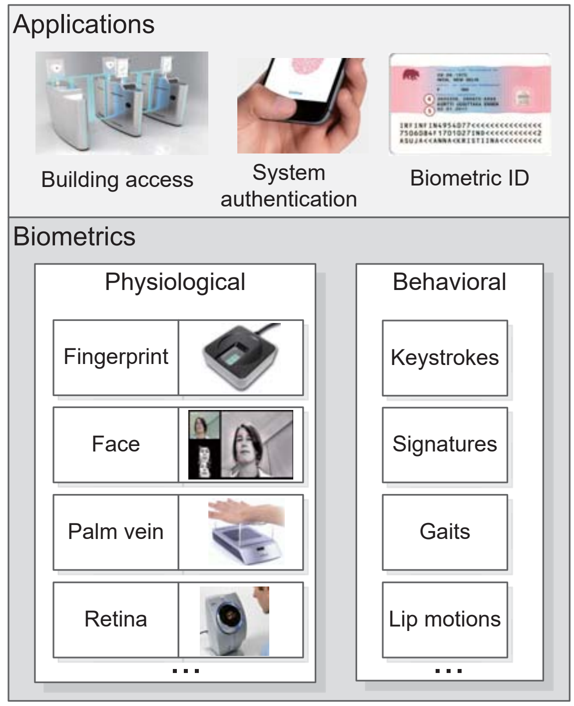
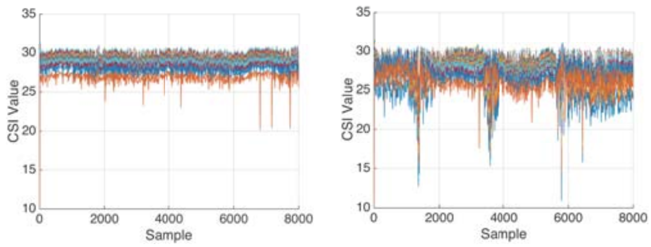


Fig. 1. The category and applications of biometrics.

B. Channel state information and its application

Modern Wi-Fi standards like IEEE 802.11n/ac employ MIMO and OFDM to improve the link capacity of the wireless communication [12]–[14]. Channel state information (CSI) is a measure of the channel frequency response (CFR) on the subcarriers [2]. With N_{tx} transmitter antennas, N_{rx} receiver antennas and N_s OFDM subcarriers, there will be $N_{tx} \times N_{rx} \times N_s$ subcarriers for the signal transmission at the same time. These subcarriers arrive at the receiver antennas in different paths. The different paths lead to signal scatters and fading in different levels. These signals are then merged into the received signals. As a result, the different real-world objects that are traversed by the subcarrier signals can have significant impact on the received signals, which can be reversely exploited to infer the positions, movements or shapes of the objects.

CSI applications. CSI has been applied in several research topics such as indoor localization, activity detection, etc [15]. Domenico et al. [16] exploited CSI for crowd counting in indoor environment. The differential CSI features for crowd counting are compared and selected. Gaebel et al. [17] adopted CSI in authentication for augmented reality systems. Wang et al. [18] utilized CSI for human activity recognition. Wang et al. [19] used CSI to achieve the real time fall detection. The basic idea for these works is that different moving patterns of human bodies can have different impact on CSI features.



(a) CSI values without movements. (b) CSI values with periodic arm-waving.

Fig. 2. Experiment on the CSI traces with periodic arm-waving.

The difference can be amplified large enough to be used for recognition of different movements and activities. It has been confirmed by several works that centimeter-level movement can be detected by the CSI variations. Considering that some biometrics are defined by the movement of certain body parts, an intuition is that it can be used to identify the human biometrics in a wireless manner. For example, one's gait can be detected by the CSI. Since different people may have different walking rates, thigh swing amplitude and body shapes, the person can be identified with the unique features conveyed in the CSI trace.

It can also be inferred that only the biometrics that involve motions can be directly captured by wireless CSI variations. Physiological biometrics such as fingerprint, retina and palm vein. cannot be directly captured by the wireless signals. In the next subsection, we will introduce the recent attempts of using CSI to establish wireless biometric IDs.

C. *Wireless Biometric Identification with CSI*

There are some recent works that focus on using CSI to recognize certain behavioral biometrics.

1) Gaits: Gaits are a distinctive feature for individuals [20]. Several recent works have been done to exploit CSI to achieve the gaits identification. The basic idea is to extract the unique influences on the CSIs by the gaits from different people. Zhang et al. [4] used FFT-based continuous wavelet transformation (CWT) method for extracting the signals in the different frequency band. Then an efficient algorithm based on the time and frequency domain is proposed for feature selection. The work can identify one person out of a group of 2-6 persons with the accuracy up to 93%.

FreeSense [8] is the state-of-the-art work, in which principle component analysis (PCA), discrete wavelet transform and dynamic time warping techniques are combined to achieve CSI waveform-based gait identification. The segmentation is subtly designed to obtain the line-of-sight waveform from CSI time series. Extensive experiments have been conducted to confirm the performance of FreeSense. A quite high accuracy (88.9%-94.5%) is achieved in identifying individuals out of a group of 2-6 people.

2) Gestures: Gestures are another typical biometric that can be used for human identification. Tan et al. [21] proposed Wi-Finger, which takes advantages of the fine-grained CSI available from commodity Wi-Fi devices and the prevalence of Wi-Fi network infrastructures. An environmental noise removal mechanism is proposed to migrate the impact of the environmental noise. Besides, the intrinsic gesture behavior is also detected which could be used to further distinguish different individuals. Wi-Finger can work with non-line-of-sight scenarios and Wi-Fi beacons.

Li et al. [7] exploited the CSI traces to recognize the gestures for wireless input method. The proposed scheme is based on the key intuition that the fingers of a user move uniquely and a unique pattern in the time series of CSI values is generated, based on which the gestures can be identified. The system achieves high classification accuracy for recognizing 9 digits finger-grained gestures from American Sign Language. Gestures can be used for identification in two ways. First, when performing the same gesture, different people have distinct therbligs which can be used to identify users. The other way is to associate people with different gestures, which can directly indicate specific users once the gesture is recognized.

III. A General Framework for Wireless Biometrics Based on CSI

In this section, we present a general framework for wireless biometrics identification based on CSI.

Before presentation of the framework, we first conduct an experiment to confirm and illustrate the ability of CSI to trace the movement of the real-world objects as follows. One Wi-Fi Access Point (AP) and one laptop are used in this experiment, where the laptop keeps pinging and receiving feedback from the AP. The Wi-Fi AP runs IEEE 802.11n and the laptop client is equipped with Intel 5300 NIC. One volunteer stands in the room and waves his arm periodically. No other people or objects are moving during the experiment. Figure 2 shows the CSI traces before/after the arm-waving activity from the volunteer. The traces are collected at the laptop side in sampling rate of 1000 samples per second. We can see that there are periodic rises in CSI value in Figure

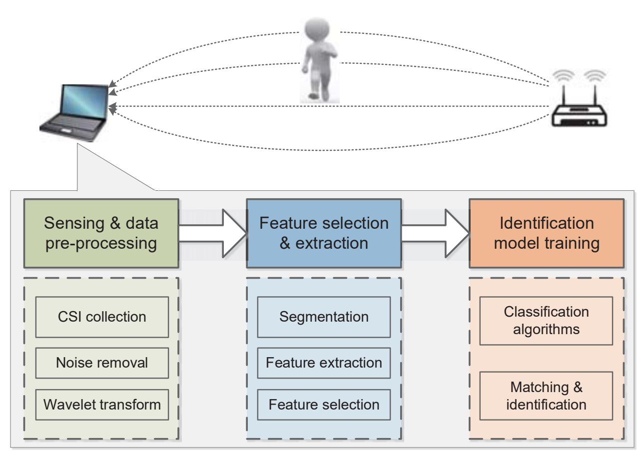


Fig. 3. The general framework of CSI-based wireless biometric identification.

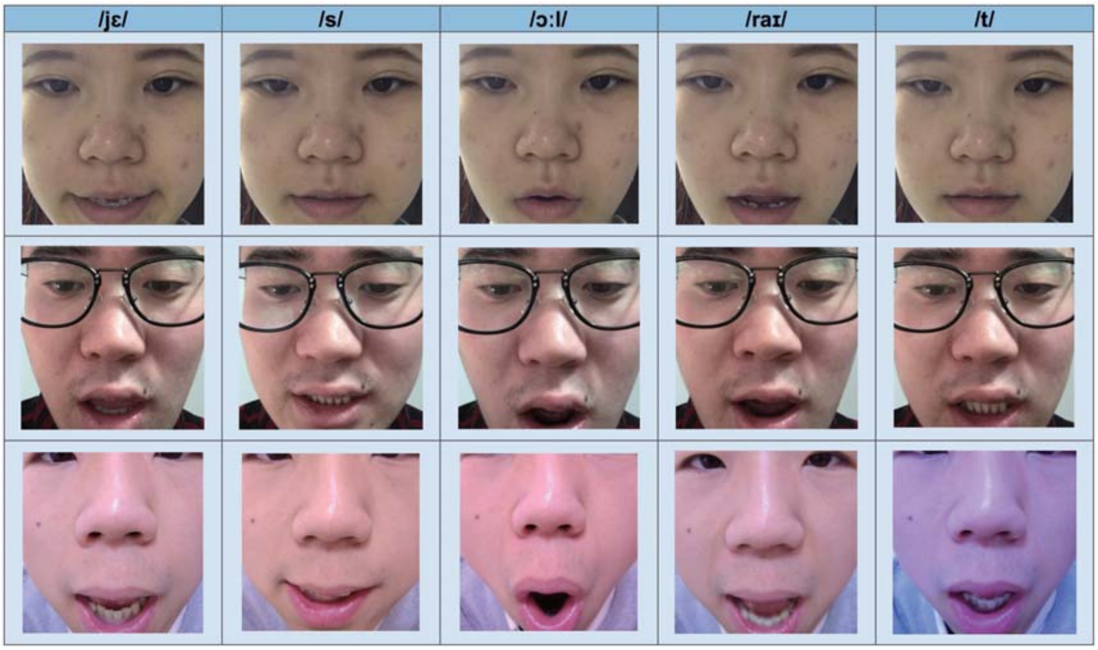


Fig. 4. The lip motions of the three volunteers. Each volunteers pronounce "yes, alright". Each row shows the lip motions of one volunteer and the five sub-figures are the captures for "ye", "s", "al", "righ", "t", respectively.

2(b) other than in Figure 2(a). Clearly these rises are caused by the periodic arm waving since it is the only difference between the two experiments. Then it can be inferred that the periodic CSI variations can be used as an indication to identify the arm-waving activity. It is also worth noting that recognizing the biometrics in real-world environments is much more challenging than that in the above experimental environment. The challenges and schemes will be discussed in the following part of this article.

Figure 3 shows the proposed framework. One Wi-Fi AP and a receiving client is required in the same area with the target user(s). The system framework consists of sensing & data preprocessing phase, feature extraction & selection phase and identification phase. In the sensing and data preprocessing phase, CSI traces are collected and the irrelevant data and noise is filtered out. The wavelet packet decomposition is used to transform the signals to time and frequency domain, which can be further used to enhance the multi-scale data analysis. In the feature extraction & selection phase, segmentation is to divide the signals into segmentations corresponding to the therbligs motions such as though swings for gaits and lip lifting for speaking. Then different algorithms can be used to extract the feature extraction and selection such as principle component analysis. Finally in the identification phase, certain training model designs are used to classify and identify the biometrics of different people. The identification results can then be further fed into the classifier to improve the identification accuracy.

For each of the three components in the framework, different algorithms and appropriate designs can be used for different biometrics. The existing works that utilize gaits and gestures can be included in the framework. Their difference is that the features and algorithms are different. The main challenges in choosing the algorithms and features for the three modules in the framework include: 1) Noise removal. Due to the large number of subcarriers, the movement of any objects can impact the received CSI significantly. Therefore, it is a non-trivial task to remove the noise data from the background interference. The key to overcome this problem is to utilize the unique time/frequency signal features to filter out the irrelevant signals. 2) Feature extraction. Different biometrics have different impact on the wireless signals. The impact on signals may not be so clear and distinct to be obtained. To overcome this challenge, it is important to combine appropriate feature extraction schemes and the biology background to find the CSI-level features.

IV. Wireless Biometric ID Based on Lip Motions: A Case Study

In this section, we follow the proposed framework and utilize the lip movements as the biometric ID for human identification. Compared to gaits and gestures, lip motions are more suitable for scenarios without much movement but with more talks such as in meetingrooms or building access authentication. We require different people speak the same sentences, and then identify specific users based on their difference in lip motions.

Different people have different talking behaviors, resulting in different lip moving rates and amplitude [22]. Considering the mouth shapes are inherently different for different people, the difference in rates and amplitude is further amplified. We conduct an experiment in which three volunteers pronounce the same words "yes, alright". Figure 4 shows the camera pictures of the three volunteers' lip motions. It can be seen that volunteer 1's amplitude is small during the pronunciation, while volunteer 2 and volunteer 3's lip amplitudes are larger. Volunteer 3's lower lip has a horizontal drift in pronouncing /s/ and /t/ compared to volunteer 2. The lip sizes and face shapes are also different for all three volunteers. Besides, the lengths of the pronunciation on the same syllables from different people are also different, resulting in the different lip moving frequency. These differences in amplitude and frequency domain on lips are the basis for CSI-based identification.

Following the proposed framework, the bioID based on lip motions works in three phases as follows.

A. Data pre-processing

To collect high-resolution CSI trace for lip motions, the transmitter keeps transmitting packets to the receiver and the CSI data is continuously stored at the receiver. Obviously not all data are associated with the lip motions and the *effective data* includes only the CSI data that captures the lip motions. In the application level, we can first locate the lip positions to reduce the irrelevant multipath effect with lip motions. This can be done by requiring users to speak at certain positions or similar to [23], which uses the MIMO beamforming technique to locate and focus on the mouth. The target users are assumed not moving during the speaking.

To further reduce the noise in signal processing level, we have following tasks in the data pre-processing. First, we need to localize the starting point of the user speaking such that the features will be extracted in the same time domain. With a wrong start point, the effective and distinctive signals may be dismissed. As the CSI data does not demonstrate clear amplification when lip motions start, we need to find a way to identify the start point instead of setting a threshold. Second, we need to remove 1) the background noise data (collected before and after the sentence speaking) and 2) the CSI data from the non-distinctive subcarriers. With the two steps, the effective CSI data for lip motions can be obtained.

We first come to the second problem of noise removal. Unlike gaits or gestures, there is no clear middle point for the whole duration of lip motions since speaking different words have no clear patterns as gaits. According to the biology background, the rate of lip motions is between 2-5 Hz [24]. The frequency of variations caused due to the lip movements lie at the low end of the spectrum while the noise frequency lies at the high end of the spectrum. As a result, we choose the Butterworth low-pass filter to remove the noise data without distorting the phase information in the lip movement signal. We apply Butterworth low-pass filter to all CSI series of all subcarriers for all antenna pairs. After the filtering, the most part of the high-frequency noise can be removed. We collect CSI data with a rate of 2000 samples per second and set the cut-off frequency of Butterworth filter at $f_c = \frac{2\pi * f}{s} = \frac{2\pi * 80}{2000} = 0.25$ rad/s, where f is the frequency of variations in CSI time series for lip motions and s is the sampling rate of CSI data. Figure 5 shows the result, where Figure 5(a) is the original CSI wave and Figure 5(b) is the CSI wave result from Butterworth filter. We can see that the Butterworth filter can successfully remove most of the noise from the CSI waves.

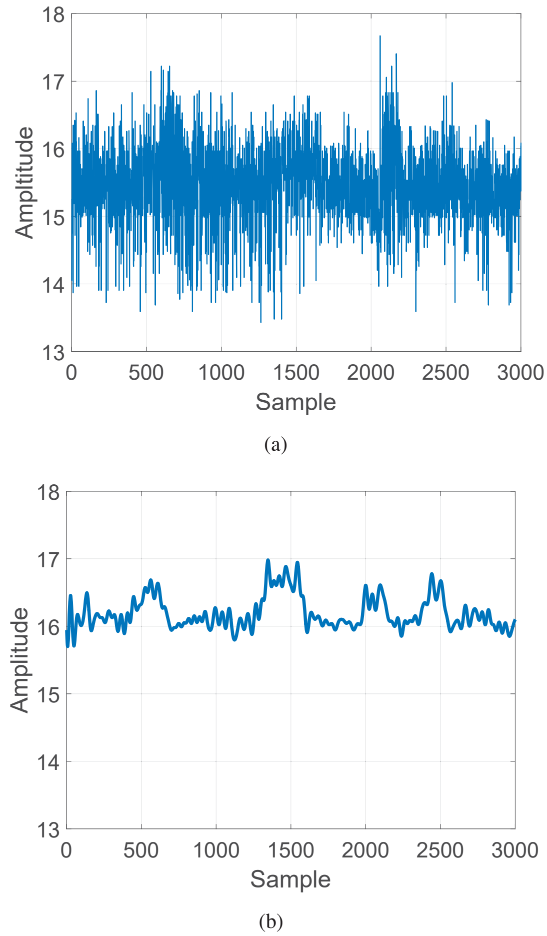


Fig. 5. The CSI wave before and after butterworth filtering.

Next, similar to the work [25], we employ PCA to further remove the CSI noise. It has been observed by several works that the changes by human activity on different subcarriers are correlated. Figure 6 shows the original CSI data for different subcarriers. We can see that though the waves of these subcarriers are quite different, there exists a strong correlation among them (e.g., the data annotated in the dashed rectangles), which is caused by the lip motions. Such correlation is quite useful for extracting the CSI data associated with the lip motions as part with the most common variations with other subcarriers is the duration affected by the lip motions. To extract the effective CSI, we use Principal component analysis (PCA) to find the principle components in CSI trace that represent the most common variations in all subcarriers, such that the effective CSI data is obtained and the noise (with uncorrelated variations) is filtered out. Besides, PCA reduces the unnecessary dimensions and can reduce the complexity in further steps. It is worth noting that the noise is hard to removed by the low-pass filter because they are quite close to the lip-sensitive CSI traces. We remove such noise with PCA due to that they are not correlated, which means they must not represent the lip motions.

With the above effective CSI data, we now solve the first problem, i.e., localizing the starting point of the lip

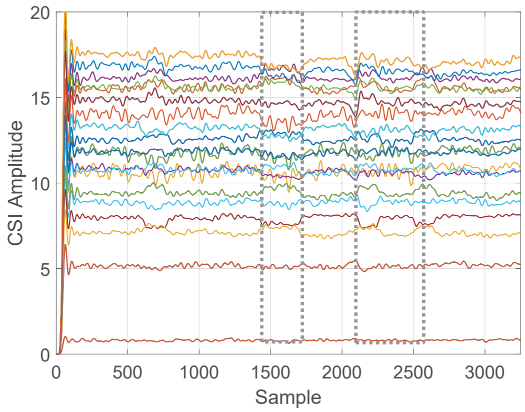


Fig. 6. The correlated variations on the CSI data of different subcarriers.

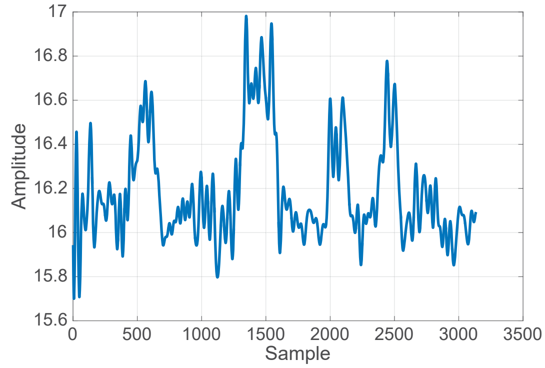


Fig. 7. The filtered CSI data for a certain subcarrier.

movements. For some users with large mouth amplitude for speaking, the starting point of the speaking can be easily obtained using a threshold τ. Figure 7 shows the filtered CSI data (of one subcarrier) for the speaking activity of a certain volunteer, whose mouth amplitude is relatively large. We can see that there are several peaks associated with several words speaking. Therefore, we can set a threshold to detect the starting point. For users with small mouth amplitude, the threshold may not work well since the micro movements are not clear enough. To this end, we exploit a two-thresholds based starting/ending point identification scheme, relying on the assumption that each word speaking will demonstrate a variation on the CSI changing rate. Instead of setting a threshold for the rate changes, similar to [25], we check the changes on the median absolute deviation (MAD) value. If the difference between the MAD values (δMAD) for two adjacent time windows exceeds a given threshold (Threshold 1), the corresponding time windows are potential starting and ending points.

For these potential starting and ending points, we calculate the short time energy e. The energy is then compared to a energy threshold (Threshold 2). If the energy exceeds the threshold, a starting or an ending point is obtained. It is worth noting that, the effective CSI data lying between the starting point and corresponding ending point is quite different from that between an ending point and its next adjacent starting point as it is expected to be silence period. The two thresholds are experimentally determined.

B. Feature extraction and selection

Now we have obtained the effective CSI data that can reflect the lip movements. To differentiate the lip motions of the same sentence from different people, we need to extract the unique features representing the the lip motions. We first test several popular features including median amplitude, standard deviation, max amplitude, max amplitude, skewness, kurtosis, dominant frequency, spectral entropy, spectral centroid, spectral flatness, spectral rolloff, zero crossing and Interquartile range. We collect the CSI data for five different volunteers and calculate the above features. The results are shown in Table I. We can see that most features are not effective

TABLE I
THE DIFFERENT FEATURES FOR THE FIVE VOLUNTEERS.

Features	User 1	User 2	User 3	User 4	User 5
median amplitude	0.2313	0.1973	0.3731	0.3846	0.0004
standard deviation	1.1742	1.0939	1.2927	1.4487	1.2314
max	2.9255	3.1468	3.0091	3.2842	3.759
min	-2.0945	-2.6711	-3.2093	-2.6907	-2.344
skewness	0.0543	-0.2492	0.1269	-0.0294	0.9305
kurtosis	2.4444	2.8842	3.1724	2.2663	3.8411
dominant frequency	0.0039	0.0117	0.0078	0.0078	0.0039
spectral entropy	0.3515	0.333	0.2264	0.0204	0.1634
spectral centroid	0.0951	0.0452	0.0638	0.2193	0.0701
spectral flatness	0.004	0.0002	0.0009	0.0005	0.0014
spectral rolloff	0.0502	0.0497	0.0449	0.0428	0.0479
zero crossing	0.0272	0.029	0.0184	0.0204	0.027
Interquartile range	1.6936	1.4342	1.2955	2.0725	1.3455

to distinguish different volunteer speakers. For example, the median amplitude values of User 3 and User 4 are close. Similarly, the kurtosis values of User 1, 2 and 4 are close. Although some features seems to be potential for user differentiation, they happen to be similar with other volunteers (not shown due to the page limit). This similarity makes it hard to distinguish different users with these features. The reason behind is that these values are essentially affected by the lip shapes and the "speaking" characteristics are not well presented with these CSI-amplitude-sensitive features.

Instead of using the above separate features, we use the extracted waveforms for the lip motions as the feature since both time and frequency information are included in the shapes. Figure 6 also shows that although the shapes are correlated, the waveforms are different. To reduce the computational complexity, we use Discrete Wavelet Transform (DWT) to compress the extracted waveforms. Compared with Short Time Fourier Transform (STFT) [26], DWT is more effective in reserving highly varying signals such as impulse and peaks. For each word speaking, we perform DWT n times. The parameter n poses a tradeoff between accuracy and computational efficiency as a larger n will lead to decreased accuracy but increased efficiency and a smaller n will lead to increased accuracy but decreased efficiency. We set a threshold for the accuracy such that the largest n value that achieves accuracy higher than the threshold is chosen. Figure 8 shows the result shape based on DWT.

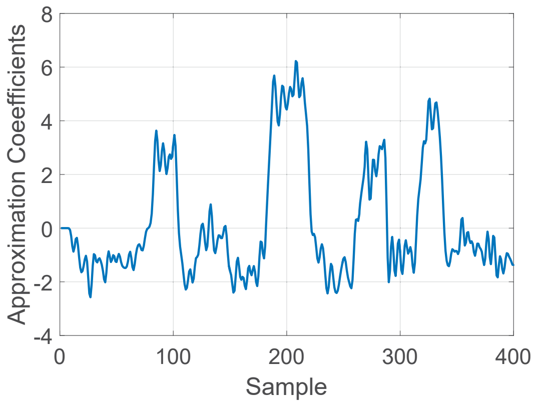


Fig. 8. The shape after discrete wavelet transform.

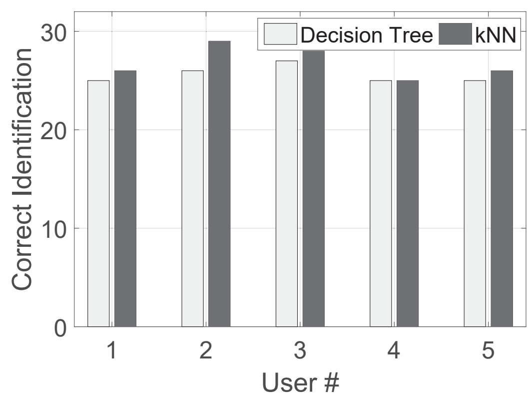


Fig. 9. The comparison between the identification with lip-based BioID based on decision tree and kNN.

C. Training and classification

When the waveform shapes are obtained via DWT, we now need to choose appropriate classifiers for biometric identification. For BioID with lip motions, we build a training model for classification with the DWT based shapes. The key problem left is how to quantify the difference among different waveform shapes, such that we can feed the distance between different shapes into existing training models. We use dynamic time wrapping to calculate the difference between different waveform shapes, which is the Euclidean distance of the optimal warping path between two waveforms calculated under boundary conditions and local path constraints. The above feature is then fed into a kNN model or a decision tree for classifying different users. We will study the accuracy of different models experimentally and discuss the reasons behind in Section V.

V. Evaluation of BioID based on lip motions

We implement BioID based on lip motions in real hardware. Specifically, we use a stationary PC equipped with Intel 5300 NIC as the CSI collector. The 5300 NIC has three antennas. The PC has an Intel i3 CPU and 16GB RAM. The operating system is Ubuntu 14.04. We use a WiFi AP with two antennas at 2.4GHz in 802.11n mode. The CSI is collected using the CSI tool for 802.11n [5]. The PC with 5300 NIC keeps pinging the AP and extracts CSI data from the feedback packets from the AP. The PC is place three meters away from the AP, and the volunteers sit in the line-of-sight between the AP and PC. No other people are moving during the experiment. We use the sentence "Hello, please open the door for me!" as the sentence for identification. Each volunteer speaks the sentence for sixty times and the CSI traces are collected, where the first half data (thirty samples) is used as the training set and the other half data is used as testing set. In order to reduce irrelevant interference, the speakers are required to lie in the office chair and keep still during the speaking.

TABLE II
The confusion matrix for the identification with lip-based BioID for the five users.

Hits	*User 1*	*User 2*	*User 3*	*User 4*	*User 5*
User 1	26	0	1	2	1
User 2	0	29	0	0	1
User 3	1	0	28	1	0
User 4	1	0	1	25	3
User 5	1	0	1	2	26

Table II shows the confusion matrix for the five users' identification. The number in each box means the number of hits. For example, the 26 in the box User1-User1 means in the 30 tests, User 1 is identified as User 1 for 26 times. From the table, we can see that 1) the average accuracy for all users is around 90%. 2) the accuracy for different users is different. The reason is that the patterns in time and frequency for their lip motions are different. User 2 and User 3 who have clearer speaking patterns (larger mouth amplitude and more unique speaking rate) achieve more accurate identification. 3) User 4 and User 5 have mutual identification errors, i.e., User 4 is identified as User 5 twice and User 5 is identified as User three times. The reason is that these two users have more similar speaking behaviors than other. However, we notice that most identification tests for those two users are still successful (86.7% and 83.3%).

We use different classification schemes and compare the accuracy. Figure 9 shows the result of identification based on kNN and decision tree for all volunteers. We can see that kNN generally achieves higher accuracy than decision tree. The reason is that the feature we used (the DWT shapes) are naturally clustered for the same person, which is more suitable for the kNN approach. We also observe that User 2 achieves the highest accuracy with kNN while User 3 achieves the highest accuracy with decision tree.

Next, we change the length of the speaking (the number of syllables) and study the accuracy for user identification. To conduct this experiment, each volunteer speaks “Hello, please open the door, OK?”. We slice the CSI data for the sentence and use part of the data set as input to study the impact of sentence length. For example, length two is used for the experiment with “Hello”. Figure 10 shows the accuracy changes for all users. It can be observed that 1) as the length increases, the accuracy for all users increases. It can be easily understood because longer speaking provides more feature data such that the difference among users can be captured. 2) The speed of the accuracy increasing becomes slower

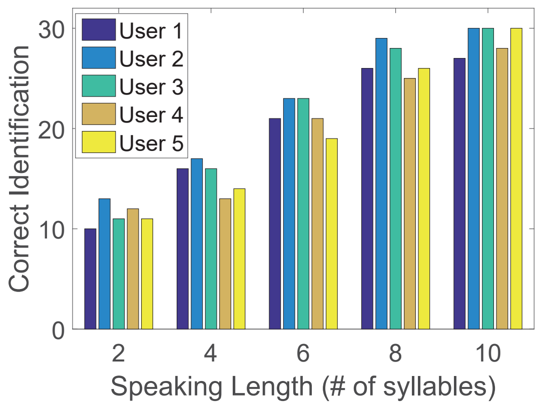


Fig. 10. The accuracy with different sentence lengths.

as the length increases. The reason is that the number of syllables increases linearly, however, its contribution to user differentiation becomes smaller as the shape distance used in kNN is already clear with several syllables.

Other possible wireless biometrics. From the above design and the cases introduced in Section II, we can see that as long as the behaviors of the biometric can *uniquely* impact the wireless signals, it can be used for human identification. Therefore, signatures, keystrokes and other biometrics with movements can also be used as wireless biometrics. For example, each person has a unique keystroke behavior, which may have unique influence on wireless signals. Therefore, even without the input words, the keystroke manner can be used to identify individuals. The implementation of the keystrokes/signature based identification can also follow the framework presented in Figure 3. The key is to design appropriate data preprocessing and feature extraction schemes based on the biological background of the keystrokes/signatures.

VI. Conclusion

The massive CSI data available on commercial-off-the-shelf NICs has led to evolutions from the specialized human-identification devices to wireless signal based recognition software. Wireless biometric IDs are one of the CSI-based evolutionary applications. In this paper, a general recognition framework of CSI-based wireless biometric identification is proposed. Based on the framework, a novel lip motion based scheme and the performance evaluation are devised. By carefully incorporating the CSI characteristics of lip motions to the noise removal and feature extraction, the proposed scheme can accurately identify different users. We conduct experiments on Intel 5300 NIC and the results show that the user identification achieves high accuracy (90%).

References

[1] A. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*. Springer Science & Business Media, 2007.

[2] Z. Yang, Z. Zhou, and Y. Liu, "From rssi to csi: Indoor localization via channel response," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, p. 25, 2013.

[3] 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.

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

[5] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011.

[6] 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*. ACM, 2015, pp. 53–64.

[7] H. Li, W. Yang, J. Wang, Y. Xu, and L. Huang, "Wifinger: talk to your smart devices with finger-grained gesture," in *Proceedings of ACM UbiComp*, 2016, pp. 250–261.

[8] T. Xin, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "Freesense: Indoor human identification with wi-fi signals," in *Global Communications Conference (GLOBECOM), 2016 IEEE*. IEEE, 2016, pp. 1–7.

[9] M. Choraś, "The lip as a biometric," *Pattern Analysis and Applications*, vol. 13, no. 1, pp. 105–112, 2010.

[10] E. Gomez, C. M. Travieso, J. Briceno, and M. Ferrer, "Biometric identification system by lip shape," in *Security Technology, 2002. Proceedings. 36th Annual 2002 International Carnahan Conference on*. IEEE, 2002, pp. 39–42.

[11] U. Park, H.-C. Choi, A. K. Jain, and S.-W. Lee, "Face tracking and recognition at a distance: A coaxial and concentric ptz camera system," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 10, pp. 1665–1677, 2013.

[12] S. N. Nazar, R. L. Olesen, P. Marinier, G. Pelletier, K. J.-L. Pan, M. Rudolf, M. U. Fazili, and L. R. Iyer, "Channel state information transmission for multiple carriers," Jul. 12 2016, uS Patent 9,391,736.

[13] G. Wang, Q. Liu, R. He, F. Gao, and C. Tellambura, "Acquisition of channel state information in heterogeneous cloud radio access networks: Challenges and research directions," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 100–107, 2015.

[14] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "Pilot: Passive device-free indoor localization using channel state information," in *Distributed computing systems (ICDCS), 2013 IEEE 33rd international conference on*. IEEE, 2013, pp. 236–245.

[15] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "Csi-based indoor localization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300–1309, 2013.

[16] S. Di Domenico, M. De Sanctis, E. Cianca, and G. Bianchi, "A trained-once crowd counting method using differential wifi channel state information," in *Proceedings of the 3rd International on Workshop on Physical Analytics*, 2016, pp. 37–42.

[17] E. Gaebel, N. Zhang, W. Lou, and Y. T. Hou, "Looks good to me: Authentication for augmented reality," in *Proceedings of the 6th International Workshop on Trustworthy Embedded Devices*, 2016, pp. 57–67.

[18] 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 ACM MobiCom*, 2015, pp. 65–76.

[19] 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*, 2016.

[20] J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S.-M. Makela, and H. Ailisto, "Identifying users of portable devices from gait pattern with accelerometers," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, vol. 2. IEEE, 2005, pp. ii–973.

[21] S. Tan and J. Yang, "Wifinger: leveraging commodity wifi for fine-grained finger gesture recognition," in *Proceedings of ACM MobiHoc*, 2016, pp. 201–210.

[22] Y. J. Meng, L. X. Bai, W. J. Liu, and M. W. Liu, "Extraction algorithm of lip characteristic parameters based on interpolation," in *Applied Mechanics and Materials*, vol. 536, 2014, pp. 235–240.

[23] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with wi-fi!" *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2907–2920, 2016.

[24] J. R. Williams, "Guidelines for the use of multimedia in instruction," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 42, no. 20. SAGE Publications Sage CA: Los Angeles, CA, 1998, pp. 1447–1451.

[25] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using wifi signals," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015, pp. 90–102.

[26] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.