Poster: EchoAuth: Gait-based Smart Home Intrusion Detection through Acoustic Sensing

Changlai Du, Xu Yuan
University of Louisiana at Lafayette
Lafayette, LA, USA
{changlai.du, xu.yuan}@louisiana.edu

***Abstract*—In this paper, we propose EchoAuth, an intrusion detection system for smart homes by taking advantage of speakers and microphones on commodity devices to capture fine-grained human gait patterns. Specifically, our system re-uses the speakers to transmit high frequency acoustic signals and the microphones to listens the reflected signals. As the reflected signals have unique variations in their Doppler shifts due to the differences in gaits from different people, they enable us to use signal processing techniques to generate the spectrograms to profile human gait patterns. By extracting effective features (that can reflect different human patterns) from the Doppler profiles, a machine learning based method is proposed to detect the intrusion for smart homes. We implement the EchoAuth system and conduct experiments by collecting a dataset with 180 gait instances from 9 people. The experimental results show that EchoAuth can achieve 93% intrusion detection rate.**

I. Introduction

Home security systems are important to families by protecting their homes and properties, and keeping the families safe from potential break-ins. However, only 27% of households in the US will adopt home security systems[1]. Cost is the most commonly cited reason for canceling professionally monitored security services. On the other hand, self-installed systems will be in high-demand because of their low cost, especially in the proliferation of Internet of Things (IoT) devices and the advance of Artificial Intelligence (AI). One example of reusing these existing sensors is the speakers and microphones widely equipped in voice assistant systems, which have been playing important roles as human helper by detecting voice commands so as to make timely response.

We propose EchoAuth as a human authentication system which uses speakers and microphones on commodity devices to capture fine-grained human gait patterns. Specifically, EchoAuth re-uses the speakers to transmit high frequency acoustic signals and listens to the reflected signals on the microphones. The reflected signals have variations in their Doppler shifts that represents the motion of the nearby human. The pattern of human limb-movement while walking is typical to the person, which is defined as the person's gait pattern. By characterizing the velocities of these movements, the Doppler shift represents a signature that is specific to the person. We then use signal processing techniques to generate the spectrograms to profile human movements. A set of human gait parameters are extracted first for the purpose of user identification. Because of the low accuracy and distinguishability of these parameters, we further propose a new set of features from the spectrogram. A machine learning based method is used to construct a multi-class identifier and spoof detector. Compared to other gait recognition systems, EchoAuth is significantly light-weight, easy to deploy and non-intrusive. EchoAuth is implemented as an app which does not require extra permissions except the access to speakers and microphones. EchoAuth reuses the always-on microphones of voice assistants and does not cost extra battery energy. EchoAuth is non-intrusive, which does not require to deploy any specific hardware such as floor sensors or wearable sensors. Unlike WiFi based methods [2] which cannot be deployed to commodity devices because the required channel state information is not directly available, EchoAuth is ready to be deployed on commodity devices.

The workflow of EchoAuth using a single speaker-microphone pair is depicted in Fig. 1. The speaker continuously transmits acoustic signals in the near-ultrasound band at 20khz and tracks the reflected signals using the microphone. EchoAuth extracts the human gait pattern signatures from the signal spectrograms which represents the Doppler profiles due to human movements. In the training phase, EchoAuth stores the extracted profiles in a local database. In the monitoring phase, it compares the detected profile with those stored in the database. The machine learning based classifier then makes the decision which class the detected profile is mostly fit.

We implemented EchoAuth on several android phones. We evaluate EchoAuth in typical smart homes. We collect a dataset with 180 gait instances from 9 people. Experimental results show that EchoAuth can achieve 93% intrusion detection rate.

II. System Design

A. Walking Detection

The first step towards identifying a person's gait is to determine which time slice of data should be used for gait extraction. To eliminate the influence of environmental noise, we keep tracking the noise level threshold N_t. N_t is updated as following:

$$N_t = (1-\alpha)N_{t-1} + \alpha E_t \tag{1}$$

where N_{t-1} is the noise level threshold at time $t-1$, E_t is the current environment noise level at time t, and α is a parameter and set to 0.1 in our settings. Human motion is detected when the signal level around 20kHz is over the noise level threshold.

B. Interference Cancellation

Since the smartphone speaker and microphone are omni-directional, signals received by the microphone include not

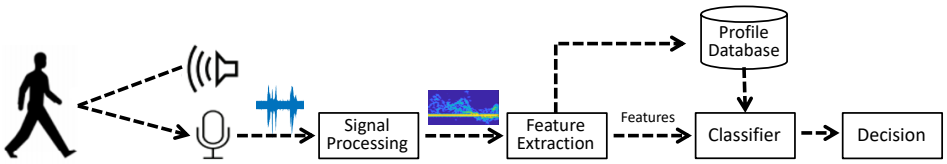


Fig. 1: The workflow of EchoAuth using a single pair of speaker/microphone.

only desired reflection from the target, but also two types of interference: (i) direct transmission and (ii) background reflection. These signals overlap in time. Even worse, the direct transmission is 3 - 4 orders of magnitude larger than the target reflection. To minimize the impact of interference, we leverage the fact that direct transmission and background reflections do not generate Doppler shift as we assume 1) the speaker and microphone are on the same device so there is no relative movement between them; 2) the device stays static so there is also no relative movement between the device and the static environment. These assumptions guarantee that we can safely delete the strong peak at 20kHz. For a setting of 10Hz frequency resolution, we reset the power value between 19.98kHz and 20.02kHz to a minimum value. We leave 20Hz on each side of 20kHz as the buffer band.

C. Parameters Measurement

The human walking model of Boulic is used, which describe the human motion with three parameters. These motion parameters can be estimated from the temporal maximum, minimum and centre velocity of the human motion distribution.

We estimate *gait cycle time*, which is defined as the time duration between two consecutive events that the right heel touches the ground. The contours of the Doppler spectrogram are used for the estimation. We estimate the *speed of human torso and human legs*. The spectrogram shows dominant torso reflections characterised by the average torso speed, torso cycle frequency and torso deviation. *Velocity* is estimated using the Doppler shift frequency.

D. Feature Extraction

The previously measured metrics (such as walking speed and gait cycle time) can be used to characterize human walking patterns. However, obtaining these measurements from a spectrogram is challenging. In order to improve the classification accuracy, we further extract features directly from spectrograms. The amount of energy associated with a particular frequency (relative to the carrier frequency) can be associated with human body moving with a certain velocity at a particular time. We therefore extract simple acoustic measurements that would capture the distribution in spectral energy as a function of time. We identify the contours of the spectrogram as well as frequency band energies and energy band frequencies[3] as the new features.

E. Intrusion Detection using Machine Learning

Every entry in training data set is labeled as either home owner or intruder. Once the gait model is trained, the classifier

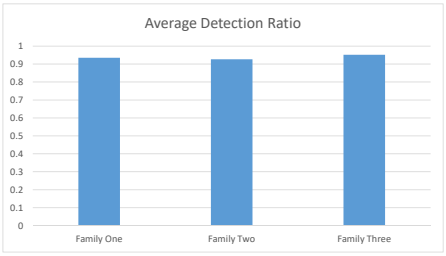


Fig. 2: Error rates in three families.

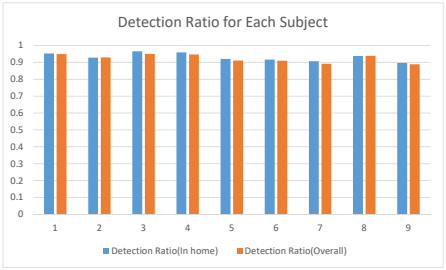


Fig. 3: Average accuracy for the nine subjects in their own home and overall.

can calculate the fitness probability that an unknown gait instance belongs to the intruders. We treat gait instances with fitness probability higher than a given threshold as instances belonging to intruders.

III. Performance Evaluation

EchoAuth is implemented on several Android phones at 20kHz. We collected data in three families with 9 adults in total. Fig. 2 shows that the detection accuracy is in comparable level across three families. The average intrusion detection rate of the three families are 93%. We also use the data set for user identification. Fig. 3 shows the detection accuracy of each subject.

References

[1] Parks associates predicts about 27 percent of u.s. households to have security by 2021. http://www.securitysystemsnews.com/article/parks-associates-predicts-about-27-percent-us-households-have-security-2021. Accessed on 2018-12-15.

[2] Wei Wang, Alex X Liu, and Muhammad Shahzad. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 363–373. ACM, 2016.

[3] Linghan Zhang, Sheng Tan, and Jie Yang. Hearing your voice is not enough: An articulatory gesture based liveness detection for voice authentication. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 57–71. ACM, 2017.

EchoAuth: Gait-based Smart Home Intrusion Detection through Acoustic Sensing

Changlai Du and Xu Yuan
University of Louisiana at Lafayette, Lafayette, LA, USA

Background and Motivation

Background:

- Home security systems protect families from potential break-ins.
- 73% of households in the U.S have no home security system because of high cost of commercial monitoring systems [1].

Motivation:

- Voice assistant systems are prevalent in houses.
- Sensing capability of voice assistant system can be used for home security purpose.
- The "always-on" acoustic signal can be used to detect unsolicited invasion with low cost.

Goal: Leveraging speakers and microphones equipped in voice assistant systems to capture fine-grained human gait patterns for intrusion detection.

Feasibility Study

- Steps observed through acoustic signals follow the similar patterns of alternating peak and valley with accelerometers.
- The difference between the shape of steps and step lengths from two persons are quite clear.
- For any given people, the acoustic-based gait remains similar between each test round in 2 hours.

EchoAuth

- EchoAuth is a gait-based smart home intrusion detection system.
- Acoustic signals are used to detect the gait patterns based on the signal process and Doppler effect.
- Machine learning based method (SVM) is applied to extract gait features from Doppler profiles.

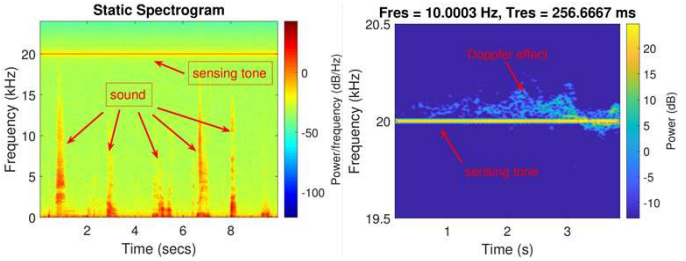
System Design

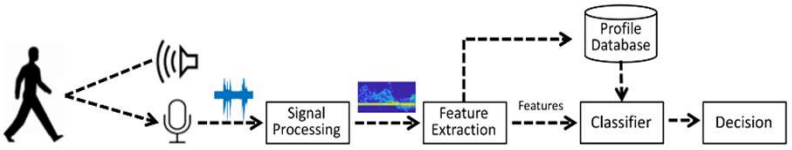
Signal generation:

- Use speakers and microphones on voice assistant system.
- Transmit "always-on" acoustic signals in the near-ultrasound band at 20kH.

Signal Processing:

- Filter and normalize time domain signals.
- Calculate the signal spectrograms.





Feature extraction:

- Extract three gait parameters: gait cycle time, velocity, speed of human torso and human legs.
- Extract detailed features from spectrograms: contours of the spectrogram, frequency band energy levels, energy band frequencies [2, 3].

Training:

- Record the extracted profiles of family members.

Detection:

- Extract gait features of the intrusion.
- Employ machine learning method (SVM) to classify different persons.

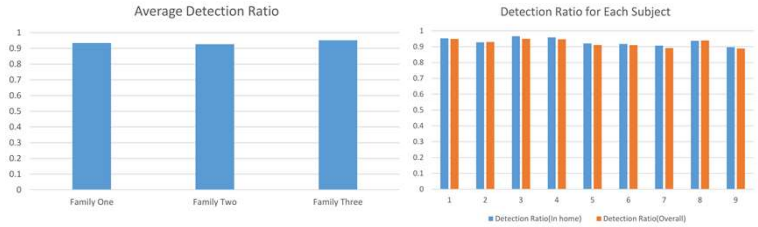
Performance Evaluation

System Setup:

- Implemented on a LG Nexus 5 and Samsung Note 5.
- Data collected from 9 adults of 3 families, with 180 gait instances in total.

Intrusion detection accuracy:

- 93% average successful rate of the three families



User identification accuracy:

- Over 90% identification rate for most subjects

Effective distance:

- Detect human walking over 5 meters
- Extract gait information within 3 meters

References

[1] Parks associates predicts about 27 percent of U.S. households to have security by 2021. http://www.securitysystemsnews.com/article/parks-associates-predicts-about-27-percent-us-households-have-security-2021. Accessed on 2018-12-15.

[2] Wei Wang, Alex X Liu, and Muhammad Shahzad. Gait recognition using wifi signals. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pages 363–373. ACM, 2016.

[3] Linghan Zhang, Sheng Tan, and Jie Yang. Hearing your voice is not enough: An articulatory gesture based liveness detection for voice authentication. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pages 57–71. ACM, 2017.