Touch-And-Guard: Secure Pairing Through Hand Resonance

Wei Wang[†‡], **Lin Yang**[‡], **Qian Zhang**[‡]
[†] School of Electronic Information and Communications, Huazhong University of Science and Technology
[‡] Hong Kong University of Science and Technology
{gswwang,lyangab,qianzh}@cse.ust.hk

ABSTRACT

Securely pairing wearables with another device is the key to many promising applications, such as mobile payment, sensitive data transfer and secure interactions with smart home devices. This paper presents *Touch-And-Guard (TAG)*, a system that uses hand touch as an intuitive manner to establish a secure connection between a wristband wearable and the touched device. It generates secret bits from hand resonant properties, which are obtained using accelerometers and vibration motors. The extracted secret bits are used by both sides to authenticate each other and then communicate confidentially. The ubiquity of accelerometers and motors presents an immediate market for our system. We demonstrate the feasibility of our system using an experimental prototype and conduct experiments involving 12 participants with 1440 trials. The results indicate that we can generate secret bits at a rate of 7.84 bit/s, which is 58% faster than conventional text input PIN authentication. We also show that our system is resistant to acoustic eavesdroppers in proximity.

ACM Classification Keywords

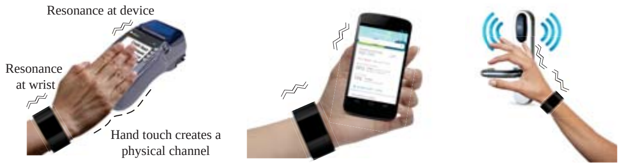
K.6.5 Management of Computing and Information Systems: Security and Protection; H.5.2 Information Interfaces and Presentation: Interaction styles

Author Keywords

Secure piaring; modal analysis; resonance; wearable

INTRODUCTION

Interacting with devices in proximity is becoming an intrinsic feature of today's wearables. This need stems from many innovative applications that provide unobtrusive experience to users. Examples are mobile payment [1] that allows users to make purchases by simply putting their phones or wearables near a contactless reader; wireless data transfer [14] that uploads health and fitness data sampled by wearables to nearby smartphones or data hubs; and smart lock [3] that un/locks wireless-chip-equipped doors by sending commands via nearby smartphones and wearables.



(a) Mobile payment. (b) Sensitive data transfer. (c) Smart home interactions.

Figure 1. Candidate applications of TAG. TAG facilitates wristband wearables to establish a secure link to another device through hand touch. The touch-based secure pairing is an intuitive and effective solution for mobile payment, data transfer, and smart home interactions.

These interactions normally involve sensitive information, which fuels the need for wearables to secure communication channels from malicious eavesdroppers [9]. The de-facto approach to set up a secure link between two devices is based on reciprocal information that is secretly shared by both sides. The reciprocal information can be manually entered PIN codes or pre-defined gestures [7, 31], bits generated from auxiliary channels using dedicated sensors [17], or ambient signals [28, 20]. However, the current crop of wearables lack conventional input interfaces and the special sensors required to support these secure pairing techniques. For instance, many wristband fitness trackers are only equipped with vibration motors and accelerometers, while they may not have touch screens on which to type PIN codes, light sensors [17] to capture laser signals, or even microphones [28] to record ambient sound. Moreover, ambient-signal-based approaches [28, 20] heavily rely on the continuous existence of ambient signals, and are vulnerable to nearby eavesdroppers.

In this paper, we show that hand touch can be used as an auxiliary channel to securely pair wristband wearables and touched devices in an intuitive manner, as illustrated in Figure 1. We design *Touch-And-Guard (TAG)*, a system that generates shared secret bits from hand touch using vibration motors and accelerometers, which are equipped in almost all smartphones, smartwatches, and wristband fitness trackers. Our observation is that the hand and the touched device form a vibration system whose resonant properties can be measured by the accelerometers in both devices. In contrast, proximate eavesdroppers can barely learn the resonant properties without physically touching the hand-device system. The resonant properties of the system is highly sensitive to different hands, devices, and how the hand touches the device. Consequently, a rich context of touch postures, positions and hand differences among users [32] leads to different resonant properties, thereby providing

UbiComp '16, September 12-16, 2016, Heidelberg, Germany
© 2016 ACM. ISBN 978-1-4503-4461-6/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2971648.2971688

enough randomness to generate secret bits. The design of TAG is inspired by modal analysis [12] in mechanical engineering. Modal analysis determines the structural vibration properties of an object by exciting it with forces of different frequencies. To extract common information from the vibration responses, we model our system as a vibration system and analyze resonant properties shared by both sides. Then, we carefully design an encoding scheme to extract secret bits from the shared resonant properties.

To validate our system, we conduct a series of experiments with 12 study participants and 1440 trials. In our experiments, each participant wears a wristband equipped with an accelerometer and touches an object attached with an accelerometer and a vibration motor. We test our system with various touch gestures, locations of the wristband, and objects of different materials. The results show that we can generate 13.72 secret bits on average in 1.75 seconds for each touch trial. The amount of secret information generated per touch is comparable to a 4-digit Bluetooth PIN code (13.2 bits). The bit rate is 7.84 bit/s, which is 58% faster than the conventional PIN input [7]. The average bit mismatch rate is merely 0.467%, and the success rate of pairing is 96%, which demonstrates the robustness of our system. Through empirical study, we demonstrate that our system is resistant to microphone eavesdroppers at various distances.

The main contributions of this work are summarized as follows.

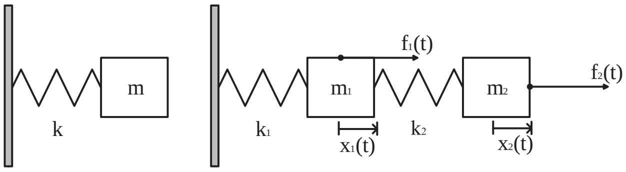
- We develop TAG, a new and intuitive way to securely pair wristband wearables with nearby devices. To the best of our knowledge, we are the first to leverage resonant properties for secure pairing.
- We propose an algorithm to extract reciprocal information from hand resonance using a haptic vibration motor and accelerometers. The ubiquity of vibration motors and accelerometers in today's wearables and mobile devices presents an immediate market for the proposed touch-based secure pairing approach.
- We test our system on 12 participants with 1440 trials in total, and conduct extensive experiments under various conditions. The results show that we can generate secret bits at a speed of 7.84 bit/s and achieve 96% success rate in establishing a secure channel. Additionally, we empirically demonstrate that acoustic eavesdroppers in proximity can learn little information about the generated bits.

CHARACTERIZING HAND RESONANCE

Modal Analysis

A mechanical system's resonant properties are determined by its physical characteristics, including its mass, stiffness and damping. A principal method to analyze the mechanical properties of a system is to break it down into a set of connected elements.

In most cases, a mechanical system is modeled as a complex multi-degree-of-freedom (MDoF) system, whose physical characteristics are represented as matrices. A complex MDoF system can be represented as the linear superposition of a number of single degree-of-freedom (SDoF) characteristics. For simplicity, we illustrate the mechanical properties using an SDoF model. As presented by Figure 2(a), an element can be characterized by an infinitely rigid constant mass m with elasticity represented by an ideal massless spring of constant stiffness k.



(a) A system of a single element.

(b) A system of two elements in physical contact.

Figure 2. A simplified model of the TAG system. A device or hand can be modeled as a single element. A hand and a touched device can be modeled as two coupled elements.

In the TAG system, the hand and the device can be modeled as two elements. When the hand touches the device, the system can be modeled as two elements with interactions, as depicted in Figure 2(b). When external forces $f_1(t), f_2(t)$ are applied to the system, the dynamic response of the system is governed by the following equation.

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{f}, \tag{1}$$

where $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$ is the displacement vector, and the $\ddot{\mathbf{x}}$ is the second-order derivative of $\mathbf{x}$. $\mathbf{M} = \left[\begin{smallmatrix} m_1 & 0 \\ 0 & m_2 \end{smallmatrix}\right]$ is the mass matrix, $\mathbf{K} = \left[\begin{smallmatrix} k_1+k_1 & -k_2 \\ -k_2 & k_2 \end{smallmatrix}\right]$ the stiffness matrix, and $\mathbf{f} = \left[\begin{smallmatrix} f_1 \\ f_2 \end{smallmatrix}\right]$ the force vector.

The displacements can be written in the form of Fourier transforms:

$$x_n(t) = \sum_{\omega} X_n(\omega) e^{i\omega t}, n = 1, 2, \tag{2}$$

where $X_n(\omega)$ is the Fourier coefficient of $x_n(t)$. In our system, the vibration motor with frequency ω_0 can be expressed in the form of a δ function as follows.

$$f_n(t) = F_n \sum_{\omega} e^{i\omega t} \delta(\omega - \omega_n), n = 1, 2, \tag{3}$$

where the δ function is defined by

$$\delta(\omega - \omega_n) = \begin{cases} 1, \omega_n = \omega \\ 0, \omega_n \neq \omega. \end{cases} \tag{4}$$

Taking (2) and (3) into (1), we yield

$$(\mathbf{K} - \omega^2 \mathbf{M}) \begin{bmatrix} \sum_{\omega} X_1(\omega) \\ \sum_{\omega} X_2(\omega) \end{bmatrix} = \begin{bmatrix} F_1 \sum_{\omega} \delta(\omega - \omega_1) \\ F_2 \sum_{\omega} \delta(\omega - \omega_2) \end{bmatrix}. \tag{5}$$

Based on (5), we can derive the frequency response function (FRF) of each element in the system, which describes magnification factors under the forces of different frequencies. The magnification factor is defined to be the ratio of the steady-state displacement response amplitude to the static displacement. As the closed-form expression is quite complex, we illustrate the resonance properties using a concrete example. We set $k_1 = 6, k_2 = 3, m_1 = 2, m_2 = 1, f_1(t) = 0$, and $f_2(t)$ can be any single frequency force. The FRF is depicted in Figure 3.

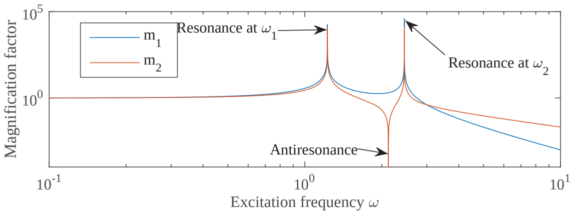


Figure 3. An illustration of resonance properties in a system of two coupled elements. We set $k_1 = 6, k_2 = 3, m_1 = 2, m_2 = 1, f_1(t) = 0$.

The FRF plot provides the following observations.

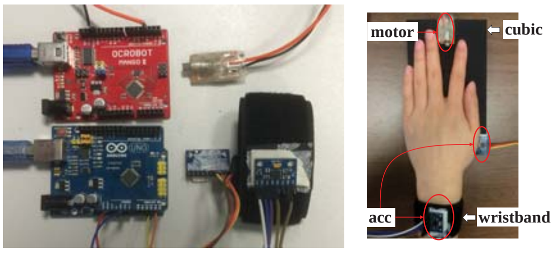
- First, the resonance properties of the two elements are consistent. Specifically, the resonant frequencies of the two elements are completely aligned with each other, and the antiresonant frequency of one element is roughly aligned with the local minimum frequency of the other element.
- Second, there are as many resonant frequencies as the number of DoFs in the system. Note that although we only model one object as a SDoF with one element, the actual object is a MDoF system consisting of multiple elements. In practice, there are many resonant frequencies in the hand-device system.

These observations imply that the resonance properties can be used as reciprocal information to generate enough secret bits for secure pairing.

Feasibility Study

To validate the above observations, we designed a prototype as shown in Figure 4. The prototype consists of a wristband with a triple-axis accelerometer, a cubic with a haptic vibration motor and a triple-axis accelerometer. We use the InvenSense MPU-6050 sensors as the accelerometers, which are equipped in many commercial wearables and smartphones. The sampling rate of accelerometers is 250 Hz. We use an Eccentric Rotating Mass (ERM) motor, which is widely adopted in today's smartphones. We use an Arduino development board [2] to control the motor to to sweep from 20 Hz to 125 Hz.

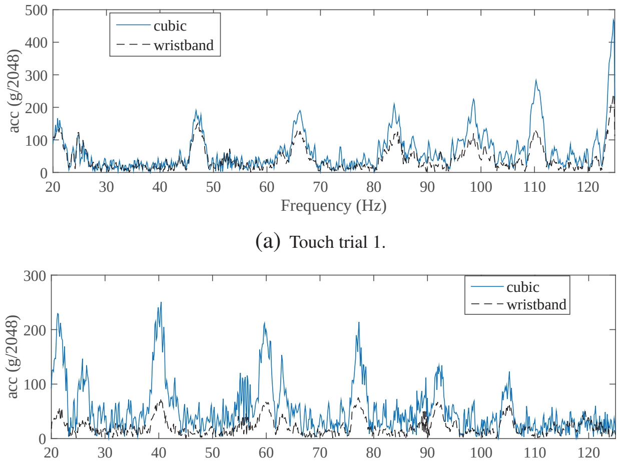
We ask participants to touch the cubic with the hand wearing the wristband as depicted in Figure 4(b), and in the meantime the vibration motor generates sweep excitation signals. The accelerometer data at both sensors are recorded and compared. Figure 5 illustrates the fast Fourier transform (FFT) of accelerometer amplitudes in two touch trials. For both touch trials, we observe that the resonant frequencies of the cubic and the wrist are well aligned. Additionally, we observe that the resonant frequencies in the two touch trials are different. Note that these two touch trials are performed by the same person, while the touch postures, strengths, and touch positions are slightly different. It indicates that resonant frequencies are quite sensitive to how a user touches an object, thereby making resonance a unique signature to each touch trial.



(a) The TAG Prototype.

(b) Touch test.

Figure 4. Prototype setup.



(a) Touch trial 1.

(b) Touch trial 2.

Figure 5. FFT of accelerometer data collected at the wristband and the cubic.

SYSTEM DESIGN

Overview

Figure 6 gives an overview of TAG, which extracts shared secrets from hand resonance for secure pairing. TAG considers a scenario where a user intends to establish a secure communication channel between its wearable and another device. The user triggers this pairing intent by touching the device. Then, the touched device generates vibration signals via a vibration motor. The vibration signals are designed to excite the device and the hand. As such, the accelerometers on the wristband wearable and the device can capture the vibration responses of the hand and the device, respectively. The wearable and the device separately process their own accelerometer data to extract reciprocal information without any information exchange. The accelerometer data process includes three steps: frequency response extraction, resonance encoding, and reverse channel coding. The frequency response extraction step screens out noise and disturbance caused by the environment and hand movements, and derives the desired frequency responses for resonance analysis. After obtaining the frequency responses, resonance and antiresonant frequencies are identified and encoded in the resonance encoding step. The reverse channel coding aims to reduce the discrepancies between the encoded bits by the wearable and the device. In particular, the original encoded bit sequences are considered as messages with a limited number of errors, and are converted into shorter sequences using a error correction code (FEC) decoder. The output of the reverse channel coding is the reciprocal information shared by the wearable and the device. It is worthwhile noting that in a complete secret sharing protocol, information reconciliation

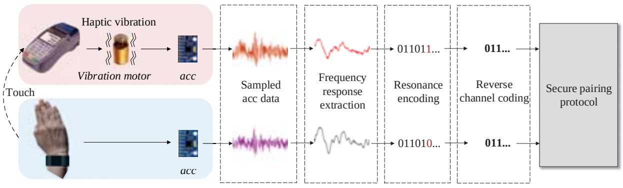


Figure 6. An overview of the TAG system.

and privacy amplification are performed to extract more reliable secrets. The reciprocal information is used to establish a secure channel. After successful pairing, the wearable notifies the user by a specific haptic feedback.

Vibration Excitation

TAG is inspired by modal analysis in that resonance properties can be derived by exciting the target object with forces of different frequencies. To this end, TAG utilizes an ERM vibration motor as the excitation source. ERM vibration motors are widely equipped in today's mobile devices to provide haptic feedback and vibration notifications. The motors are supplied with DC power and rotate an eccentric mass around an axis to create a centripetal force, which causes the motors and the attached devices to vibrate. The centripetal force is the external force applied to the hand-device system, and can be expressed as

$$f(t) = md\omega^2 \sin(\omega t), \tag{6}$$

where m is the eccentric mass, d the distance from the center of gravity to the center of rotation, and ω the angular velocity of the rotation. The motors tune the input voltage to control the angular velocity ω, which determines the amplitude and the frequency of the force. In practice, the analog sinusoidal waveform is approximately generated with binary voltage levels using Pulse Width Modulation (PMW). In particular, PMW modulates the duty cycles of the DC power to simulate a voltage between the DC power voltage and zero voltage. In our system, we generate the vibration excitation by controlling the duty cycles of the DC power. Specifically, we gradually increase the duty cycles to generate forces with sweeping frequencies. Figure 7 gives a visual illustration of the our vibration excitation.

The frequency range needs to be selected carefully to obtain resonance properties. Previous studies [4, 5] have reported that the natural frequencies of the human hand-arm systems range from several Hertz to hundreds of Hertz. Therefore, a subset of the resonant frequencies of the hand-device system fall within this range. Apparently, the wider frequency range we select, the more complete resonance properties we can obtain. However, the maximal frequency that can be captured by an accelerometer is gated by its sampling rate. According to the Nyquist sampling theory, a sensor at f sampling rate can capture signals at frequencies no more than $f/2$. As most of the accelerometers equipped in today's mobile devices support up to 400 Hz sampling rates, the maximal frequency is gated by 200 Hz. In addition, there is a tradeoff between the frequency range and the vibration duration. For a given frequency sweeping speed, the vibration duration is proportional to the frequency range. In our implementation, we select 20-125 Hz as the frequency range, which is managed to generate secret bits comparable to a 4-digit PIN code.

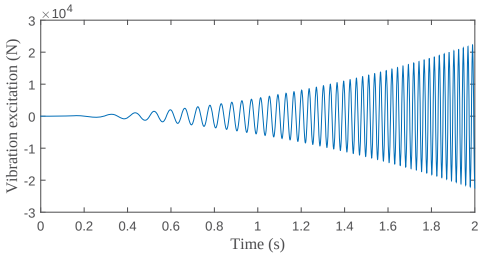


Figure 7. An illustration of vibration excitation.

The speed of frequency sweeping determines the duration of one touch trial. We aim to set the sweeping speed as fast as possible to minimize the touch duration. The limit of the sweep speed is gated by the transient state duration. When an external force changes its frequency, the forced system needs a short period of time before reaching the steady state. Note that we can only accurately obtain the resonance properties when the system is in the steady state. However, it is hard to identify which part of the accelerometer data is collected in the steady state, as the duration and patterns of the transient state depend on many confounding factors. For each vibration frequency, the collected accelerometer data contains two parts: the data in the transient state and the data in the steady state. To amortize the impact of the transient state, the vibration motor is set to stay for enough time before increasing its frequency. As such, the amount of data in the steady state is dominant and the overall data retains strong resonant properties. We empirically evaluate the system under various durations and set the motor to sweep from 20 Hz to 125 Hz within 1.75 s, which eliminates the impact of the transient state.

Frequency Response Extraction

To extract resonant properties, we first need to derive the frequency response of the hand-device system from the raw accelerometer data. We observe that the accelerometer data at low frequencies is largely polluted by motion artifacts. In practice, it is inevitable that the hand moves during the pairing process. The acceleration caused by motion is usually much larger than the vibration-induced acceleration, thereby making it hard to extract vibration-induced acceleration. Fortunately, the frequencies of motion artifacts concentrate at low frequencies of several Hertz [34, 18]. Hence, we set the minimal vibration frequency to over 20 Hz to avoid overlapping with hand motion frequencies. As such, we only extract resonant properties in the vibration frequency range where motion artifacts are negligible.

Recall that although the resonant frequencies of different elements match each other, their responses at other frequencies are not identical, as illustrated in Figure 3. These mismatches in real systems are much more complex, and lead to local variances which might mislead the resonant frequency identification. We observe that there are multiple peaks near one resonant frequency. Thus, these local variances should be mitigated before performing resonance encoding. To this end, we use a moving-average filter to eliminate these local variances. We empirically find that the smoothing window of 10 samples is enough.

Resonance Encoding

Resonance encoding translates the frequency response into a sequence of bits. After local variance removal, we obtain two highly similar curves in the frequency domain. To encode frequency responses, we have the following alternative options: 1) encoding the amplitudes of the frequency response by quantizing the amplitude of each frequency or frequency segment into multiple levels; 2) encoding the shape of the frequency response curve by classifying the curve of each frequency segment into several predefined shapes, such as ascending and descending shapes; and 3) encoding the positions of resonant and antiresonant frequencies. Although the first and second options can preserve most of the information, they are inapplicable in our case. As we observe in Figure 5, the amplitudes of the two frequency responses are not coincidental. Thus, amplitude quantization would introduce many mismatches, which would lead to a high failure rate in pairing. The shape-based encoding faces a similar issue, as the two curves do not coincide in non-resonant frequency ranges. Therefore, we turn to the third option that encodes the resonant and antiresonant frequencies to ensure the matching rate.

Our encoding algorithm consists of two steps: resonant and antiresonant frequencies identification and modulation. We use local maxima and minima in the frequency response to identify the resonant and antiresonant frequencies. We employ a sliding window to move across the whole frequency range, and find all the extrema (i.e., maxima or minima) in each sliding window. Note that there may be multiple extrema near one resonant or antiresonant frequency due to local variances. We observe that resonant frequencies separate from each other by at least 10 Hz. To avoid repetitive extrema, we select at most one maximum and minimum in each sliding window of 10 Hz. In particular, if there are multiple minima or maxima in one sliding window, we select a winner based on amplitude and discard the others. After scanning the whole frequency response, the frequencies of all extrema are marked as resonant or antiresonant frequencies.

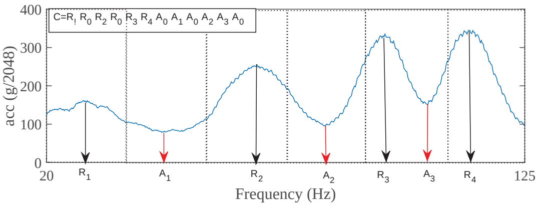


Figure 8. An illustration of resonance encoding. The relative location of each resonant or antiresonant frequency in its segment is encoded. The encoded bit sequence C consists of encoded locations of resonant frequencies $\{R_i\}$ and antiresonant frequencies $\{A_j\}$. Segments without resonance or antiresonance are encoded as R_0 or A_0.

Then, we modulate these frequency locations into a sequence of bits. An intuitive method is to quantify frequencies and encode these frequency levels. However, this encoding method leaks certain information as it has predictable patterns. The order of resonant frequencies (e.g., in an ascending or descending order) must be preset so that the two sides can derive the same sequence of bits, which leaks information in the encoded bit sequence. For example, if the resonant frequencies are encoded in an ascending order in the bit sequence, eavesdroppers know that the first codeword in the bit sequence is likely to be small as it corresponds to the minimal resonant frequency. To avoid such information leakage, we encode the relative locations rather than the absolute locations of resonant frequencies. First, we divide the whole frequency range into N segments. Then, we encode the relative locations of resonant and antiresonant frequencies in the corresponding segment that covers the frequencies, as illustrated in Figure 8. To encode relative locations in a segment, we evenly divide a segment into m subsegments, and quantify the frequency locations based on these subsegments. Segments without resonant or antiresonant frequencies are encoded as R_0 or A_0. In our implementation, we use two bits to encode the relative locations in a segment. We divide each segment into three subsegments and use “01”, “11”, and “10” to encode frequencies in these subsegments. We set R_0 and A_0 to be “00” to encode segments without resonant or antiresonant frequencies. If there are multiple resonant (or antiresonant) frequencies in one segment, we select the frequency with higher (or lower) amplitude for encoding. Empirical results show that there are 4-8 resonant (or antiresonant) frequencies in 20-125 Hz. Hence, we divide the frequency range into 6 segments.

Reverse Channel Coding

After resonant encoding, the wearable and the device derive n-bit sequences, denoted as C_w and C_d, respectively. Due to noise and mismatched local variances, C_w and C_d may differ at certain bits. To correct these error bits, we employ reverse channel coding (RCC), which trades off bit mismatch rate with bit rate. RCC aims to convert two bit sequences with slight differences into one identical codeword of shorter length. In particular, we treat C_w and C_d as inputs, and use an FEC decoder that maps C_w and C_d to their closest k-bit codewords. As C_w and C_d contain quite small numbers of different bits, they can be mapped to the same codeword with high probability.

EXPERIMENT DESIGN

Experimental Setup

To validate the TAG system, we conducted experiments using an experimental prototype as depicted in Figure 4. The prototype uses an Arduino OCROBOT Mango II development board to control an ERM vibration motor, and an Arduino UNO development board to collect acceleration data from two InvenSense MPU-6050 sensors. The vibration motor and one accelerometer is attached to an object, while the other accelerometer is worn on a wrist of the participant using a wristband. To simulate the scenario of mobile payment, we use a cubic box as the mobile payment end. The cubic size is 6.3 in (length) × 3.8 in (width) × 1.9 in (height), as shown in Figure 4(b). In addition, we also attach sensors to a smartphone, a mouse, and a cup as the touched objects.

The maximal input voltage of the ERM motor is 3.3 V. We developed an application to control the input voltage of the

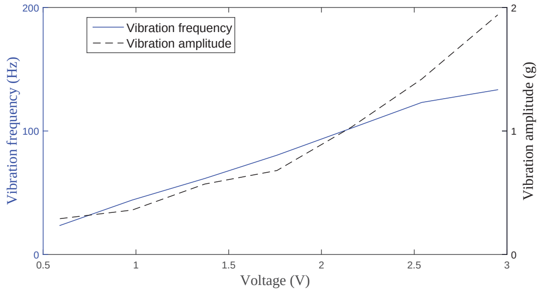


Figure 9. Specifications of our ERM motor.

motor using PWM. The vibration amplitudes and frequencies of the motor under different voltages are measured and shown in Figure 9. The sampling rate of the accelerometer sensors is set to be 250 Hz to capture all vibration responses. The acceleration data is collected via an Arduino board and processed offline using MATLAB R2014b.

We use an iPhone 5s as an acoustic eavesdropper that records vibration-induced sound through its built-in microphone. The iPhone 5s is placed in the proximity of 1-36 inches away from the motor. We use the built-in microphone to record acoustic signals during our experiments with a sampling rate of 44.1 kHz. The recorded data during each touch trial is uploaded to a PC, and is processed using the same algorithm to infer the bit sequence derived from the acceleration data. The experiment environment is in a quiet office so that vibration-induced acoustic signals are not overwhelmed by background noise. The sound pressure level (SPL) of the office during our experiments is around 40-50 dB.

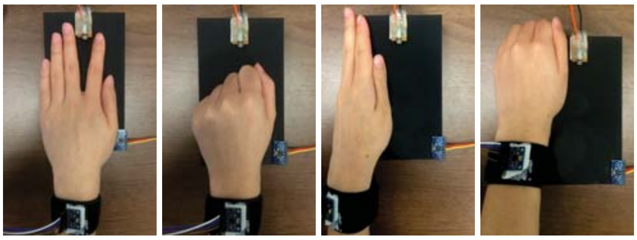
Enrolled Participants

We invite 12 volunteers, including 5 females and 7 males, with ages ranging from 23 to 31. We specifically select subjects to cover a wide range of wrist circumferences and body mass indices (BMI). Wrist circumference and BMI are import physical attributes related to hand vibrations, as our system should be robust for users of different physical attributes. In particular, the wrist circumferences range from 5.51 inches to 7.48 inches, and BMI ranges from 17.5 to 27.70.

Procedure

Prior to touch trials. The touched object was placed on a desk in our office. An iPhone 5s was placed on the same desk at a distance of 6 inches away from the object to eavesdrop acoustic signals leaked from the vibration. Note that we varied the distances in our security validation experiment. Prior to starting touch trials, we demonstrated the performance of different touch postures. We performed four touch postures, including palm touch, fist touch, border touch, and corner touch, to touch different areas of the object, as illustrated in Figure 10.

Performing touch trials. Each participants were asked to wear a wristband equipped with an accelerometer on its preferred hand, and use that hand to touch the object. Seven participants chose to wear the wristband on their left hands while five others chose to wear it on their right hands. The wearing locations of the wristband were based on the participants' own habits of wearing watches or wrist wearables. Then, each participants was asked to perform four different touch postures as we demonstrated in Figure 10. We only showed different contact areas of these touch postures without specific requirements on touch strength, or detailed hand/arm gestures. The participants were asked to repeat each touch posture 30 times. One touch trial lasted 1.75s, during which the motor vibrated with sweeping frequencies from 20 Hz to 125 Hz, while the iPhone 5s used its built-in microphone to records acoustic signals. The participants were allowed a small rest period of around 5 s between trials of a posture, and a longer break of 10-30 s between different postures. We yielded a dataset with 1440 trials, where each participant contributed 120 trials. We collected additional trials from 4 participants in controlled settings to study the impact of vibration durations and wearing locations. Each participant performed 30 trials in each vibration duration and wearing location setting.



(a) Palm touch. (b) Fist touch. (c) Border touch. (d) Corner touch.

Figure 10. Touch postures.

EVALUATION

Evaluation Metrics

We employ the following metrics to evaluate the performance of our system.

- *Bit rate*. We use bit rate to measure how fast we can generate reciprocal information from resonant properties. Given the number of secret bits (13.29 bits for a 4-digit PIN code) required for pairing, a higher bit rate indicates a shorter time needed for pairing. In our system, bit rate depends on the vibration duration and the encoding scheme. Recall that the vibration duration is gated by the period of the transient state. The vibration duration should be long enough so that the effect of the transient state does not overwhelm the resonant properties in the steady state.

- *Bit mismatch rate*. Bit mismatch rate is defined as the ratio of mismatched bits to the total number of generated bits. A lower bit mismatch rate indicates a higher probability that the wearable and the device generate the exact same sequence of bits. The bit mismatch rate is also affected by the vibration duration and the encoding scheme. There is a tradeoff between bit mismatch rate and bit rate. Longer vibration duration yields stronger resonant properties, thereby achieving lower bit mismatch rate at the cost of lower bit rate. We need to identify the optimal vibration duration that delivers the highest bit rate while maintaining strong resonant properties for encoding.

- *Entropy*. Entropy measures the average amount of information contained in a message [10]. The entropy of a random variable X is computed by $H(X) = -\sum_{i=1}^{n} \Pr[x_i] \log_2 \Pr[x_i]$, where $\Pr[x_i]$ is the probability of X's possible value x_i. In our evaluation, we compute entropy per segment to measure the uncertainty of the generated secret bits. The probability of each bit is computed by counting its frequency in repeated trials. The secret bits with higher entropy contain more information, and are harder for eavesdroppers to infer.

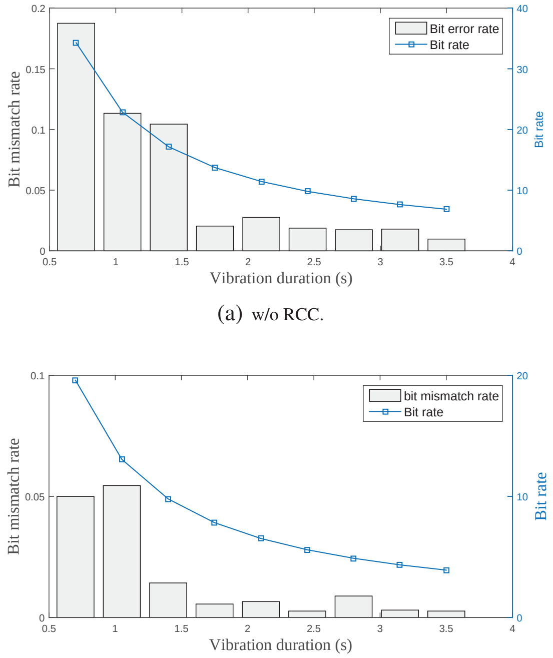
- *Mutual information*. Mutual information is a measure of the amount of information about one random variable obtained through another random variable [10]. We use mutual information to measure the information leakage in our system. Less mutual information between two random variables X and Y indicates that one can learn less about X by observing Y. Mutual information close to zero between the bit sequences obtained by the eavesdropper and those of the wearable or the device indicates that the eavesdropper is unable to obtain any useful information about the bit sequences generated from resonant properties.

Pairing Performance

This section studies the pairing performance of our system in terms of bit mismatch rate and bit rate. First, we conducted a set of micro-benchmark experiments to evaluate the impact of different settings. We varied the vibration durations to find an optimal duration for one touch trial (Figure 11). In order to test the robustness of our system, we asked participants to wear the wristband at different locations (Figure 12). Then, we followed the setup as described in the **Procedure** section and obtained the overall performance across all participants (Table 1 and Figure 13).

A key factor that affects the bit mismatch rates and bit rates is the vibration duration. We need to identify the optimal vibration duration that minimizes the negative impact of the transient state to achieve bit mismatch rate with the maximal bit rate. To this end, we empirically study the performance under various vibration durations as shown in Figure 11. The results of encoding schemes with and without RCC are illustrated. We employ Hamming(7,4) as the channel coding scheme. The results show that the bit mismatch rates diminish quickly when the vibration duration is larger than 1.5 s, while the improvement is minimal when we further extend the duration beyond 1.75 s. The results imply that the vibration duration of 1.75 s is long enough to extract reliable bits at a rate of 13.71 bit/s for the scheme without RCC and 7.84 bit/s for the scheme with RCC. The bit rate in our system outperforms that of the conventional PIN code input, whose bit rate is 4.96 bit/s, according to the experiments in [7]. In the following evaluation, we set the vibration duration to 1.75 s.

In real scenarios, the wearing locations of wrist wearables vary among users. In our experiments, the wristband is put on locations according to participants' habits of wearing watches or wearables. Before proceeding to the results under this uncontrolled wearing setting, we conduct a separate experiment in which we intentionally vary the locations of the wristband to investigate the robustness of our system. We ask participants to place the wristband close to their wrist joints (location 1), and move the wristband 0.5 inch (location 2), 1 inch (location 3), 1.5 inches (location 4), and 2 inches (location 5) away from their wrist joints. Figure 12 shows that the bit mismatch rates of the scheme without RCC increase slightly when the wearing location moves away from the wrist joint, while those of the scheme with RCC stay below 0.8% across all locations. The reason behind the results is that the vibration amplitudes of the hand resonance decay when propagating along the forearm, thereby making it harder to accurately identify the resonant and antiresonant frequencies at the wristband. Fortunately, as the bit mismatch rates are still lower than 6%, the scheme with RCC can still correct most of these errors. Moreover, the cases wearing wearables more than 1 inch away from the wrist joint are quite rare. We observe that most of the natural wearing locations of the participants fall into the range between location 1 and location 2.



(a) w/o RCC.

(b) w/ RCC.

Figure 11. Bit mismatch rates and bit rates under various vibration durations.

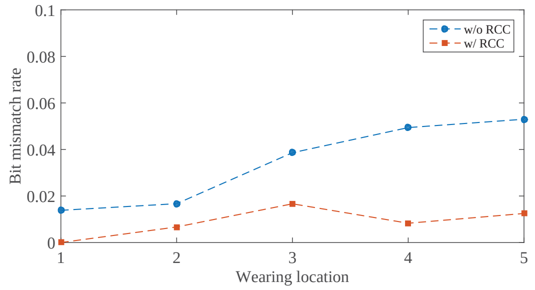


Figure 12. Bit mismatch rates with different wearing locations.

The overall performance with 1.75 s vibration duration and uncontrolled wearing locations are given in Table 1 and Figure

	Palm	Fist	Border	Corner
w/o RCC	1.13%	0.57%	2.1%	3.9%
w/ RCC	0	0	0.43%	1.47%

Table 1. Bit mismatch rates of different touch postures.

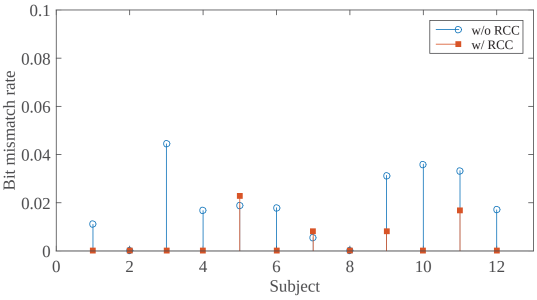


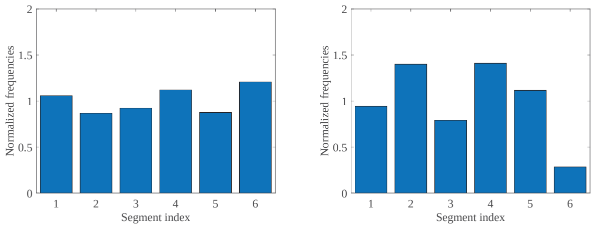
Figure 13. Bit mismatch rates of all participants.

13. The bit mismatch rates of different touch postures are summarized in Table 1. The palm and fist touch postures achieve zero bit mismatch rate under the encoding scheme with RCC, while the corner touch posture performs worst of all. The reason behind the results is that palm and fist touch postures provide larger touch areas and thus lead to stronger resonance, while the corner touch posture provides the smallest touch area. The bit mismatch rates of all touch postures are consistently low, which indicates the usability of the touch-based secure pairing. Figure 13 shows the bit mismatch rates across all participants, whose basic information is listed in Table **??**. On the whole, our system achieves bit mismatch rates of 1.93% and 0.47% for the scheme without and with the RCC, respectively. For the complete scheme, i.e., the scheme with the RCC, the successful rate of secure pairing for all trials is 96%, which indicates that generated bit sequences in 96% of the trials are completely matched. The average number of trials needed to for successful pairing is 1.04. It is worth noting that the results are comparable to other secure pairing techniques [7, 28, 25, 19].

Security Validation

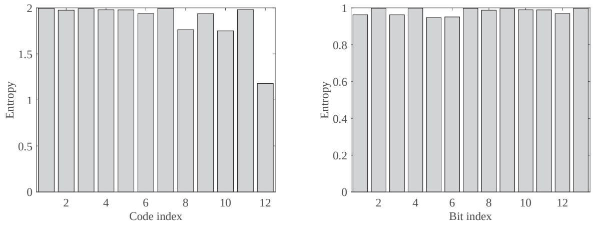
This section evaluates the security performance of our system. To ensure the reciprocal information obtained from the resonant properties is substantially unpredictable, we first measure the randomness of generated bits (Figure 14, Figure 15). Then, we study the information leakage under acoustic eavesdropping attacks (Figure 16-Figure 18).

Figure 14 measures the normalized numbers of resonant and antiresonant frequencies falling into each segment. The numbers per segment are counted based on all trials in our experiments. We observe that the normalized frequencies are comparable to each other, except for that of the antiresonant frequency in segment 6. The reason is that we may miss the antiresonant frequencies when they are near the highest vibration frequency. Nevertheless, most resonant and antiresonant frequencies are randomly distributed in different segments with comparable probabilities. This indicates that we yield reasonable randomness through our segment-based encoding scheme.



(a) Resonant frequencies per segment. (b) Antiresonant frequencies per segment.

Figure 14. Normalized numbers of resonances and antiresonances per segment.



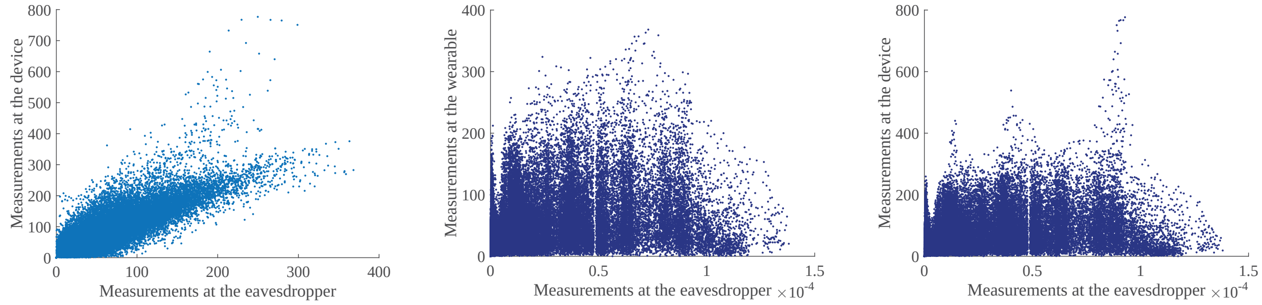
(a) Entropy per segment. (b) Entropy per bit.

Figure 15. Entropy of generated bits.

We further quantify the randomness of the reciprocal information using entropy. Figure 15(a) measures the entropy per segment of the bit sequences directly derived from the resonant properties without RCC. The code index represents each of the 12 codes that encode the resonant and antiresonant locations, as illustrated in Figure 8. We observe that the entropies of most codes approach two, which is the theoretical upper-bound. The entropy of code 12 is lower as we miss some antiresonant frequencies in the last segment due to the limitation of vibration frequency. Figure 15(b) shows the entropy per bit after applying RCC. We see that the entropies of all bits approach the theoretical upper-bound, i.e., 1, indicating high randomness of the generated bits.

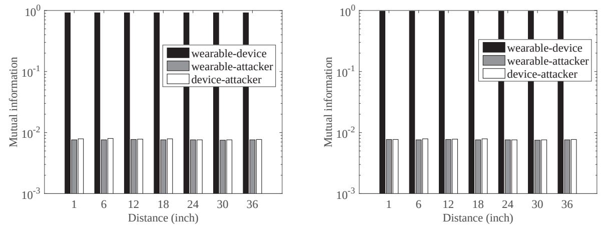
In order to evaluate the information leakage to eavesdroppers, we first compare the raw frequency responses obtained by the wearable, the device, and the eavesdropper. The eavesdropper's measurements are downsampled to match the acceleration data. Figure 16 shows pairwise scatterplots of the measurements collected by the three entities. The intuitive meaning of the visual results is that the measurements of the wearable and the device are well aligned with each other, while the eavesdropper's measurements are uncorrelated with those of the wearable or device. The fundamental reason behind the results is that the subtle vibrations of the hand and object incur extremely small sound, which is overwhelmed by surrounding noise and the acoustic signals generated from the motor.

To quantify how much information the eavesdropper can learn from its measurements, we empirically compute the mutual information between the bits derived by the three entities using the same encoding scheme. Figure 17 shows the mutual information under various eavesdropping distances. Eavesdroppers at different distances obtain a negligible amount of information about the wearable's and the device's measure-



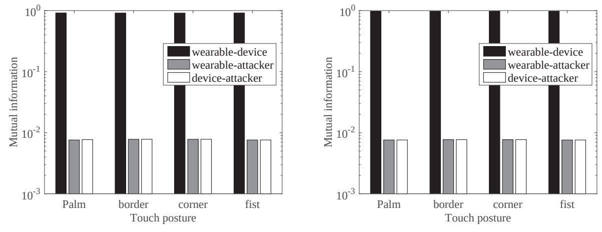
(a) Wearable vs. Device. (b) Eavesdropper vs. Wearable. (c) Eavesdropper vs. Device.

Figure 16. Comparisons of frequency response measurements. Amplitudes of each frequency is compared and plotted. We use the dataset of all trials. The eavesdropper is 6 inches away.



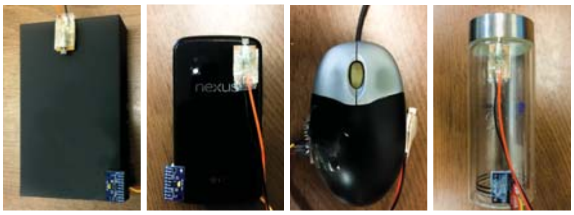
(a) w/o RCC. (b) w/ RCC.

Figure 17. Mutual information under different eavesdropping distances.



(a) w/o RCC. (b) w/ RCC.

Figure 18. Mutual information under different touch postures.



(a) Cubic box. (b) Smartphone. (c) Mouse. (d) Glass cup.

Figure 19. Different objects as the touched device.

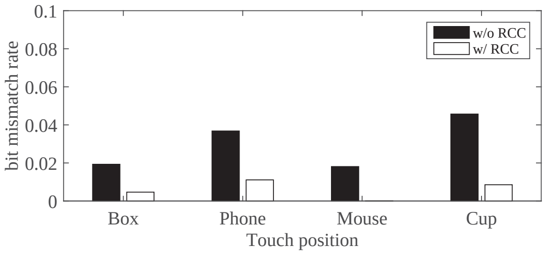


Figure 20. Bit mismatch rates of different objects.

ments. The mutual information between the eavesdropper and the wearable (device) is less than 0.01, which indicates that the eavesdropper can learn less than 0.01 bit for 1 bit of the wearable's (device's) bit sequences. Figure 18 measures the mutual information under different touch postures. The results are consistent with Figure 17, in that the eavesdropper can learn less than 1% information about the wearable's and the device's bit sequences.

Different Objects

To test the feasibility of our system on different objects, we extend our experiments by using an additional set of objects as the touched devices, as shown in Figure 19. The participants are asked to hold the smartphone or the mouse in their hands wearing the wristband. The cup is placed on the desk and the participants are asked to touch the area of the side. Other settings are the same as described in the **Procedure** section. Figure 20 shows the bit mismatch rates of different objects. The performance varies among different objects due to their different levels of resonant properties in the vibration frequency range. For all objects, the scheme without RCC achieves bit mismatch rates lower than 5%, and the scheme with RCC achieves bit mismatch rates lower than 1%, indicating the feasibility of our system on these objects.

DISCUSSION

Security Protocols

We investigated the feasibility of generating shared secret bits from the resonant properties of the hand and the touched object. It provides an intuitive means to securely pair a wristband wearable with another device. The focus of this paper is to generate shared secret bits for both sides, which is a common and essential step of most secure pairing protocols. Our system can be applied to different secure pairing protocols, including PIN-based authentication, two-factor authentication, and secret key based encryption. In particular, the secret bits generated from hand resonance can be used as the PIN code shared by both sides, a proof of physical contact for the two-factor authentication, or a basis to generate the secret key.

Accelerometer-based Eavesdroppers

The touched object in the TAG system can be held in hand or put on a desk. We evaluate acoustic eavesdropping as vibrations produce sound and may leak information to acoustic eavesdroppers in both cases. When the touched object is on a desk, accelerometer-based eavesdropping is also possible when the eavesdropper is placed on the same desk. In this case, the desk, the object, and the touched hand form a coupled system. How much information can be leaked to the eavesdropper depends on the physical properties of the desk, and the distance between eavesdropper and the object. As the mass of the desk is normally much larger than the object, resonant vibration of the desk is much weaker than that of the object and the hand. To decode the resonance-encoded information, the distance between the eavesdropper and the object needs to be quite short. The accelerometer-based eavesdropper can be easily found and thus is not a major threat to TAG.

Visual Eavesdroppers

While we have empirically demonstrated that our system is resistant to acoustic eavesdroppers in proximity, it has certain limitations. Although the subtle vibrations of hand resonance are too small to be captured by microphones, it might be recovered by high-speed cameras. A recent study [11] has successfully recovered acoustic signals from vibrations using high-speed cameras. Although the vibrations of hand resonance is weaker than audible sounds as in [11], it is still possible for a high-speed camera to recognize the subtle vibrations of hand resonance and recover the measurements of the accelerometers. Our main argument to this problem is that our system is still safe against general shoulder surfers using eyes or normal-speed cameras, which are threats to conventional PIN code methods. One simple defense to high-speed cameras is to block the line of sight. For example, we can use the other hand to cover the hand performing the pairing, much as we do to avoid shoulder surfers when typing our passwords.

RELATED WORK

Many approaches have been proposed to establish a secure link between two devices based on shared secrets. The shared secrets can be generated from user interactions, auxiliary channels, or authenticated with user actions or auxiliary channel. Examples of the former include gesture-based authentication [7, 31] that encodes authentication information as gestures defined by authenticators or users, and the techniques that require users to simultaneously provide the same drawings [29] or shaking trajectories [21]. The auxiliary channel based approaches leverage a special channel to create shared secrets. Many studies use ambient environments, such as ambient sound [28, 15], radio environment [20], or a combination of multiple environments [22] as the proof of physical proximity. The auxiliary channel itself is also leveraged as the source to generate shared secrets. Normally, the two devices send messages to each other within a short time to measure the channel between them. Qiao et al. [25] use the frequency shapes of the wireless channel between two devices to generate secret bits. Similarly, Liu et al. [19] use the channel sate information (CSI) as the shared secrets. Different from these approaches, this paper exploits a new and intuitive method that generates shared secrets through hand resonance. The advantages of our method lie in its intuitive user interaction, and the ubiquity of the required sensors, i.e., vibration motors and accelerometers, in today's wearables. Several recent advances [6, 16] have proposed to use vibration signals to generate shared secrets for physically connected devices. However, vibration signals leak over the air and can be captured by acoustic eavesdroppers.

Vibration properties of objects have been exploited to enable different applications. Ono et al. [23] develop a touch sensing technique that recognizes a rich context of touch postures based on the resonant changes when users change their touch postures and positions. SoQr [13] estimates the amount of content inside a container based on the vibration responses to acoustic excitations. VibID [32] studies the vibration properties of different persons, and design a wristband wearable to recognize household people based on their vibration properties. This paper is inspired by these studies, and takes it one step further in that we exploit the resonant properties of two objects (a hand and its touched device) in physical contact to facilitate secure pairing.

Recent advances [27, 26, 33, 8] have empowered physical vibration as a relatively slow but secure communication channel for smart devices. The vibration channel can be used to authenticate a shared key derived from a Diffie-Hellman exchange over insecure wireless interfaces. Differently, TAG aims to generate shared secrets directly from resonance and then use the secrets to secure wireless communications. However, the emitted acoustic signals cannot be completely canceled if there are multiple eavesdroppers. Similarly, intra-body communication technologies [24, 30] are developed to allow users to create a physical communication channel through hand touch. Dedicated transceivers are installed for the intra-body communications. Different from these techniques, this paper aims to generate shared secrets from the vibration channel rather than enabling a new communication channel. As such, we can truly reap the security benefits from vibration channel, while delivering high data rates from the traditional wireless communication channels such as Bluetooth or Wi-Fi channels.

CONCLUSION

This paper presents TAG, a new and intuitive approach to enable secure pairing for wearables. The insight is that a hand and its touched object form a system whose resonant properties are shared by both sides. We build a prototype to extract shared secrets from the resonant properties using commercial vibration motors and accelerometers. The ubiquity of vibration motors and accelerometers in today's smart devices maximizes the chance of widespread acceptance for our system. We demonstrate the feasibility of our system by evaluating with 12 participants. We collect 1440 trials in total and the results show that we can generate secret bits at a rate of 7.84 bit/s with merely a 0.467% bit mismatch rate.

Acknowledgment

The research was supported in part by grants from China NSFC under Grant 61502114, 61173156, 973 project 2013CB329006, ITS/143/14FP-A, RGC under the contracts CERG M-HKUST609/13 and 622613, as well as the grant from Huawei-HKUST joint lab.

REFERENCES

1. Apple Pay. `http://www.apple.com/apple-pay`
2. Arduino. `http://arduino.cc`
3. August Smart Lock. `http://august.com/products/august-smart-lock/`
4. SA Adewusi, S Rakheja, P Marcotte, and J Boutin. 2010. Vibration transmissibility characteristics of the human hand–arm system under different postures, hand forces and excitation levels. *Journal of sound and vibration* 329, 14 (2010), 2953–2971.
5. S Adewusi, M Thomas, VH Vu, and W Li. 2014. Modal parameters of the human hand-arm using finite element and operational modal analysis. *Mechanics & Industry* 15, 6 (2014), 541–549.
6. Joshua Adkins, Genevieve Flaspohler, and Prabal Dutta. 2015. Ving: Bootstrapping the Desktop Area Network with a Vibratory Ping. In *Proceedings of the 2nd International Workshop on Hot Topics in Wireless*. ACM, 21–25.
7. Imtiaj Ahmed, Yina Ye, Sourav Bhattacharya, N Asokan, Giulio Jacucci, Petteri Nurmi, and Sasu Tarkoma. 2015. Checksum gestures: continuous gestures as an out-of-band channel for secure pairing. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*. ACM, 391–401.
8. Ahmed Al-Haiqi, Mahamod Ismail, and Rosdiadee Nordin. 2014. A new sensors-based covert channel on android. *The Scientific World Journal* 2014 (2014).
9. Ming Ki Chong, Rene Mayrhofer, and Hans Gellersen. 2014. A survey of user interaction for spontaneous device association. *ACM Computing Surveys (CSUR)* 47, 1 (2014), 8:1–8:40.
10. Thomas M Cover and Joy A Thomas. 2012. *Elements of information theory*. John Wiley & Sons.
11. Abe Davis, Michael Rubinstein, Neal Wadhwa, Gautham Mysore, Fredo Durand, and William T. Freeman. 2014. The Visual Microphone: Passive Recovery of Sound from Video. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 33, 4 (2014), 79:1–79:10.
12. David J Ewins. 1984. *Modal testing: theory and practice*. Vol. 15. Research studies press Letchworth.
13. Mingming Fan and Khai N Truong. 2015. SoQr: sonically quantifying the content level inside containers. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*. ACM, 3–14.
14. Shyamnath Gollakota, Haitham Hassanieh, Benjamin Ransford, Dina Katabi, and Kevin Fu. 2011. They can hear your heartbeats: non-invasive security for implantable medical devices. *ACM SIGCOMM Computer Communication Review* 41, 4 (2011), 2–13.
15. Nikolaos Karapanos, Claudio Marforio, Claudio Soriente, and Srdjan Capkun. 2015. Sound-proof: usable two-factor authentication based on ambient sound. In *24th USENIX Security Symposium (USENIX Security)*. 483–498.
16. Younghyun Kim, Woo Suk Lee, Vijay Raghunathan, Niraj K Jha, and Anand Raghunathan. 2015. Vibration-based secure side channel for medical devices. In *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*. IEEE, 1–6.
17. Tim Kindberg and Kan Zhang. 2003. Secure spontaneous device association. In *Ubicomp*, Vol. 2864. Springer, 124–131.
18. Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter* 12, 2 (2011), 74–82.
19. Hongbo Liu, Yang Wang, Jie Yang, and Yingying Chen. 2013. Fast and practical secret key extraction by exploiting channel response. In *INFOCOM, 2013 Proceedings IEEE*. IEEE, 3048–3056.
20. Suhas Mathur, Robert Miller, Alexander Varshavsky, Wade Trappe, and Narayan Mandayam. 2011. Proximate: proximity-based secure pairing using ambient wireless signals. In *Proceedings of the 9th international conference on Mobile systems, applications, and services (MobiSys)*. ACM, 211–224.
21. Rene Mayrhofer and Hans Gellersen. 2009. Shake well before use: Intuitive and secure pairing of mobile devices. *Mobile Computing, IEEE Transactions on* 8, 6 (2009), 792–806.
22. Markus Miettinen, N Asokan, Thien Duc Nguyen, Ahmad-Reza Sadeghi, and Majid Sobhani. 2014. Context-based zero-interaction pairing and key evolution for advanced personal devices. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 880–891.
23. Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. 2013. Touch & activate: adding interactivity to existing objects using active acoustic sensing. In *Proceedings of the 26th annual ACM symposium on User interface software and technology (UIST)*. ACM, 31–40.
24. Duck Gun Park, Jin Kyung Kim, Jin Bong Sung, Jung Hwan Hwang, Chang Hee Hyung, and Sung Weon Kang. 2006. TAP: touch-and-play. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 677–680.
25. Yue Qiao, Kannan Srinivasan, and Anish Arora. 2014. Shape matters, not the size: A new approach to extract secrets from channel. In *Proceedings of the 1st ACM workshop on Hot topics in wireless (HotWireless)*. ACM, 37–42.
26. Nirupam Roy and Romit Roy Choudhury. 2016. Ripple II: Faster Communication through Physical Vibration. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. 1–14.

27. Nirupam Roy, Mahanth Gowda, and Romit Roy Choudhury. 2015. Ripple: Communicating through physical vibration. In *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. 265–278.

28. Dominik Schurmann and Stephan Sigg. 2013. Secure communication based on ambient audio. *Mobile Computing, IEEE Transactions on* 12, 2 (2013), 358–370.

29. Mohit Sethi, Markku Antikainen, and Tuomas Aura. 2014. Commitment-based device pairing with synchronized drawing. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*. IEEE, 181–189.

30. Mitsuru Shinagawa, Masaaki Fukumoto, Katsuyuki Ochiai, and Hakaru Kyuragi. 2004. A near-field-sensing transceiver for intrabody communication based on the electrooptic effect. *Instrumentation and Measurement, IEEE Transactions on* 53, 6 (2004), 1533–1538.

31. Jing Tian, Chengzhang Qu, Wenyuan Xu, and Song Wang. 2013. KinWrite: Handwriting-Based Authentication Using Kinect.. In *Proceedings of the 20th Annual Network and Distributed System Security Symposium (NDSS)*. 1–18.

32. Lin Yang, Wei Wang, and Qian Zhang. 2016. VibID: User Identification through Bio-Vibrometry. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks (IPSN)*. ACM, 1–12.

33. Takuro Yonezawa, Jin Nakazawa, and Hideyuki Tokuda. 2015. Vinteraction: Vibration-based information transfer for smart devices. In *Mobile Computing and Ubiquitous Networking (ICMU), 2015 Eighth International Conference on*. IEEE, 155–160.

34. Li Zhang, Parth H Pathak, Muchen Wu, Yixin Zhao, and Prasant Mohapatra. 2015. Accelword: Energy efficient hotword detection through accelerometer. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM, 301–315.