Identification of Pressed Keys by Time Difference of Arrivals of Mechanical Vibrations

Gerson de Souza Faria[a,1,*], Hae Yong Kim[a,1,]

[a] *Departamento de Engenharia de Sistemas Eletrônicos, Escola Politécnica, Universidade de São Paulo, Av. Prof. Luciano Gualberto, tr. 3, 158, CEP 05508-010, São Paulo, Brazil*

Abstract

The possibility of finding the sequence of pressed keys in a mechanical keyboard is a serious security threat. In our previous work, we have shown that it is possible to identify, with high probability, the pressed key by analyzing the vibration generated by the keystrokes. At that time, we did not know the physical phenomenon responsible for leaking information as mechanical vibration. In this paper, we show that the TDOA (Time Difference of Arrivals) of the mechanical waves is the main culprit for leaking information. To demonstrate this hypothesis, we glued three accelerometers in a PIN-pad, collected the vibrations generated by the keystrokes and computed the relative delays of vibration arrival times in pairs of accelerometers. We show that it is possible to estimate the positions of the keys through simple difference of the delays. A simple classification scheme using the delays yielded 96.4% of recognition success rate. The same technique can be used to attack devices with touch-sensitive screen, identifying the region touched.

1. Introduction

Mechanical keypads are widely used for entering confidential data. Confidential passwords are typed in mechanical keypads in ATMs (Automatic Teller Machines) or PIN-pads (devices used in smart card transactions to input the card-

*Corresponding author

Email addresses: gerson.faria@usp.br (Gerson de Souza Faria), hae@lps.usp.br (Hae Yong Kim)

Preprint submitted to Computers & Security *November 23, 2015*

holder's Personal Identification Number). In some countries, including Brazil, electors use electronic voting machines with mechanical keyboards to choose the candidate. Thus, the possibility that someone finds out the sequence of pressed keys, without the user's knowledge or consent, is a serious security threat. In card operations, the theft of card information in an otherwise legitimate transaction, known as "skimming", was responsible for 87% of attacks against ATMs in 2013, as reported in [1].

In a previous work [2], we have shown that it is possible to identify the pressed key with high probability by gluing accelerometers in the device, acquiring acceleration signals generated by keystrokes and analyzing these signals. We called it "vibration attack".

Usually, modern ATM keypads are encrypted. They are sealed modules that encrypt the PIN soon after the entry. So, non-encrypted PIN numbers are not meant to be accessible from outside either by physically tapping onto wires or remotely sensing electromagnetic radiation. Any tampering of the keypad causes it to permanently disable itself. Similarly, PIN-pads are protected modules that permanently disable themselves if tampered. The possibility of identifying the sequence of pressed keys through mechanical vibrations is a serious security failure of secure keypads because they are designed to resist against any attempt of eavesdropping. The devices will continue functioning normally while passwords are stolen.

When we wrote our previous paper, we did the experiments without knowing the physical phenomenon responsible for the leak of information. We extracted a lot of features from the vibration signals (up to 165 features per keystroke) and fed machine learning algorithms with them in an attempt to identify the pressed key. This was enough to certify the existence of the problem, but without a satisfactory explanation of the underlying phenomenon.

In this work, we show that the propagation delay of the transverse wave generated by the keystroke is the main phenomenon responsible for the information leaking. With this knowledge, in this work we use much less features per keystroke (2 instead of up to 165) and less training data (100 or 200 keystrokes per experiment instead of up to 2400 keystrokes) and obtain similar classification success rates than in our earlier work. This result is somewhat surprising, because PIN-pad is far from being a homogeneous medium, and one would expect that the vibration propagation velocities were different in different regions of the device. To provide our technique a short name, we will call it "vibration delay attack".

It is also possible to estimate the position of the pressed key (the source of the wave) through a simple 2-D trilateration of the relative delays of the signals captured by the accelerometers. This is a well known technique in a variety of fields by terms like TDOA (Time Difference of Arrivals) or simply "time of flight". For instance, the accurate measurement of these delays is the basis of GPS (Global Positioning System) and other geolocation systems. Geophysicists and seismologists also use it in order to locate the epicenters of earthquakes and of other seismic events [3]. In our case, the position of the key is analogous to the epicenter of an earthquake.

In the literature, there are some papers that identify the pressed key by sound, because each key usually emits a characteristic sound when pressed. Asonov and Agrawal [4] achieved 79% of key recognition success rate when identifying one out of 30 keys in a PC keyboard. Berger *et al.* [5] use keyboard acoustic emanations and a dictionary to recognize correctly 73% of the English words typed in a PC keyboard, without any training. Zhuang *et al.* [6] takes as input 10-minute sound recording of a user typing English text using a keyboard and recovers up to 96% typed characters. Halevi [7] uses keyboard acoustic emanations for eavesdropping over random passwords, without using dictionary, achieving 40% to 64% recognition rate per character.

Similarly to acoustic emission, each key seems to emit a characteristic mechanical vibration when pressed. However, this idea has been much less explored in the literature. Marquardt and Verma [8] use this idea to recognize keystrokes of a computer keyboard. They use the accelerometer of a smartphone placed near the computer's keyboard to capture the vibrations. They do not actually identify the pressed key. Instead, they classify keystrokes in "left" or "right" and pairs of keystrokes in "near" and "far". They achieved classification rates from 65% to 91% making those binary decisions.

The phenomenon identified in this work is of a different nature: even if it were possible to have all the keys emit exactly the same sound and the same mechanical vibration, it would be still possible to identify the pressed key by the arrival times of the vibration wave. Our purpose in this work is neither to select the most appropriate classifier nor to achieve extremely high recognition rates. Instead, our primary aim is to show that there is one more physical phenomenon that can be used to identify the pressed key by means of a simple location technique, but applied in a complex non-homogeneous medium. Most of location experiments use relatively homogeneous solids, like concrete, metal, glass, acrylic etc. and not composite ones, like a PIN-pad. We use in all experiments only the relative delays as features and a simple Naive Bayesian classifier. If we add other features and fine-tune the classifier, probably we would achieve higher success rates. Additionally, our finding also opens the possibility of attacking touch-screen devices, because the same phenomenon occurs when the user interacts with them. Note that touch-screen devices cannot be attacked using acoustic emanations.

The literature on trilateration comes from diverse fields of research. Maochen Ge discusses the source location theories and methods that are used for earthquake, microseismic and acoustic emission [9, 10]. He analyzes the principles of source location methods and mentions the main causes of inaccuracy, for instance, imprecision of sensor positions and errors in arrival time measuring. Geolocation methods based on measuring the time difference of arrivals (TDOAs) of signals received from several geostationary satellites are presented in [11, 12, 13]. Ho and Chan present a method that solves a set of nonlinear equations to estimate the location [11]. Gustafsson and Gunnarsson compare a Monte Carlo method and a gradient search algorithm [12]. Schumacher *et al.* propose a Bayesian approach for the problem of source location in the materials research [14]. Arun *et al.* [15] develop a location method based on Kullback-Leibler discrimination information criteria on spectra of acceleration signals, testing the method on a large aluminium plate.

The rest of the paper is organized as follows. Basic theory on transverse waves is described in Section 2. We apply the vibration delay attack in two devices: a simple mockup keypad in Section 3 and a commercial PIN-pad designed to be secure in Section 4. We make some considerations comparing the previous results with the new ones in Section 5 and present our conclusions in Section 6. Appendixes present the definition of normalized cross correlation (used to estimate the relative delay between two signals) and the source location estimation method.

2. Vibration of a Plate

2.1. Theory

The behavior of a transverse wave in a bar or plate (with thickness) is considerably more complex than the classical transverse wave in a string or membrane (with negligible thickness). Plates and bars have thickness, bringing properties as bending stiffness (also known as flexural rigidity) defined as the resistance offered by the plate while undergoing bending or deflection.

The differential equation for the deflection of a one-dimensional string is [16]:

$$\nabla^2 y(x,t) = \frac{1}{c^2}\,\frac{\partial^2 y(x,t)}{\partial t^2}, \quad c^2 = \tfrac{T}{\rho} \tag{1}$$

where T is the tension and ρ is the mass density of the material. All functions of the form $y(x,t) = F_1(x-ct) + F_2(x+ct)$, $\forall F_1, F_2$, are its solutions, where c is the constant velocity of the traveling wave without shape deformation.

On the other hand, the simplified wave equation for the transverse vibration of a uniform bar is:

$$\nabla^4 y(x,t) = -\frac{1}{a^2}\,\frac{\partial^2 y(x,t)}{\partial t^2}, \quad a^2 = \tfrac{EI}{m} \tag{2}$$

where E is the modulus of elasticity of the material, I is its moment of inertia and m its total mass. Let us assume that a solution of Eq. 2 is a simple harmonic wave traveling with velocity v:

$$y(x,t) = A\ cos\frac{2\pi}{\lambda}(x-vt) \tag{3}$$

Substituting Eq. 3 in Eq. 2, we obtain a velocity that depends on the wavelength, $v = a\frac{2\pi}{\lambda}$. Note in the previous relation that a does not possess dimensions of velocity, so it does not represent a velocity, instead of c in Eq. 1 that is in fact a velocity.

In summary, the travelling velocity of a wave is constant in a string but, in a bar, it depends on the wavelength and consequently on the oscillation frequency, because the latter is a *dispersive medium*. A sinusoidal wave can travel in a dispersive medium without suffering deformation in its shape, but a wave packet will be deformed in such a medium since its components have, by construction, distinct wavelengths. In this case, each component will travel with a distinct velocity thus causing deformation [16, 17].

The same phenomenon occurs in plates, like the acrylic plate where we made the two initial experiments (Sections 2.2 and 3).

The dispersion and reflections make it difficult to accurately measure the delays in the arrival of mechanical vibrations, because different ways of pressing keys generate distinct spectra and so different delays between wavefronts and reflection occurrences. We measure the delays of wavefronts considering them as packets travelling with a group velocity. The group velocity of a wave is the velocity with which the overall shape of the wave's amplitudes propagates.

2.2. Dispersion in Acrylic Plate

In order to observe in practice the effect of medium dispersion and group velocity presented in Section 2.1, we made an experiment in an acrylic plate using two distinct sources of excitation: (i) touching the plate with the finger and (ii) touching it with the tip of a mechanical pencil. Fig. 1 depicts the assembly of the experiment. The dimensions of the plate are approximately 3mm×640mm×670mm. We mounted the two accelerometers over small metallic screws and glued them on the acrylic plate. A_1 and A_2 are the positions of the accelerometers.

In all the experiments, we use Freescale MMA7361 analog triaxial low-g accelerometers [18] operating in ±1.5g range and a Tektronix TDS-2004B digital

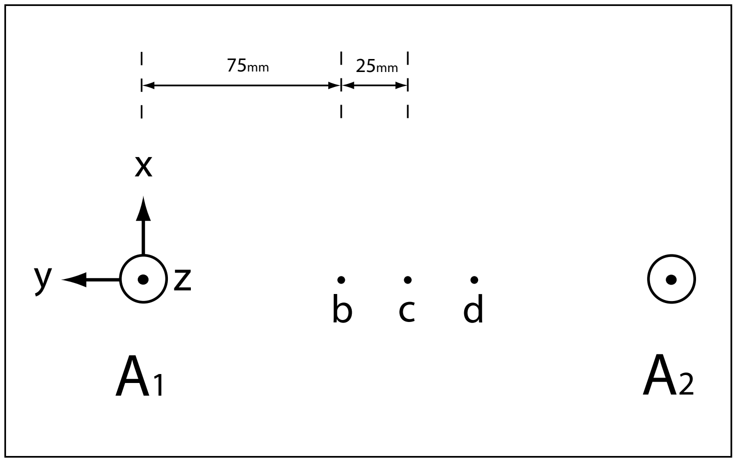

Figure 1: Assembly of the experiment to observe the medium dispersion in acrylic plate.

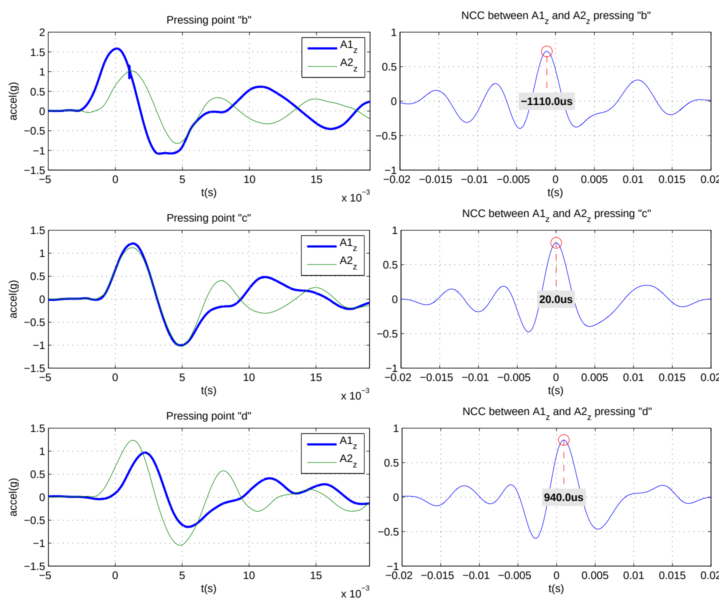

Figure 2: (Left) Acceleration signals obtained tapping the acrylic plate with finger at points ‘b’, ‘c’ and ‘d’. (Right) Relative delay estimated using the position of the highest peak in NCC between $A1_z$ and $A2_z$.

153 oscilloscope to acquire the data. Each signal vector comprises 2500 points, the
154 maximum allowed by the oscilloscope. The sampling rate varies from experiment

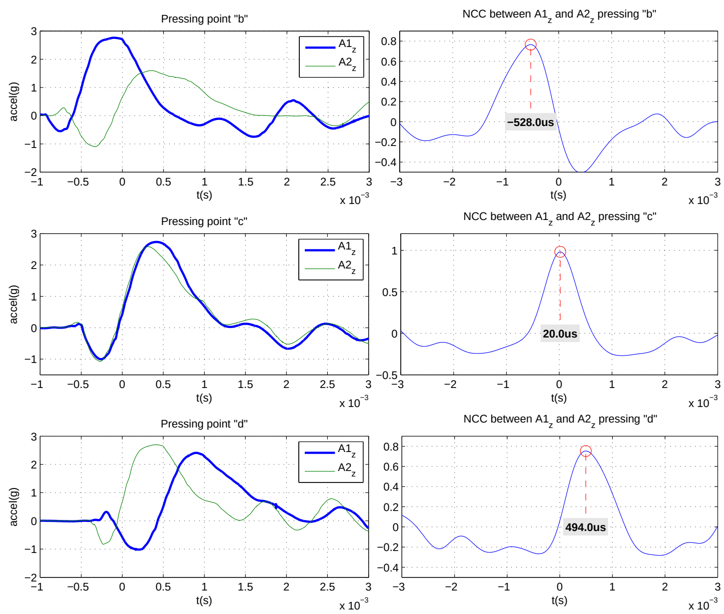

Figure 3: (Left) Acceleration signals obtained tapping the acrylic plate with a mechanical pencil at points 'b', 'c' and 'd'. (Right) Relative delay estimated using the position of the highest peak in NCC between $A1_z$ and $A2_z$.

to experiment. In this case, the sampling rate was 100KS/s for the experiment touching the plate with the finger and 500KS/s for touching it with a mechanical pencil.

The left column of Fig. 2 depicts the transverse $\vec{z}$ acceleration signals acquired by the accelerometers, touching the plate with the finger at points "b", "c" and "d". Longitudinal waves in $\vec{x}$ and $\vec{y}$ directions are much faster than transverse waves because they travel inside the material and not on its surface. So, we ignored the longitudinal signals, processing only surface transverse signals $\vec{z}$.

Obviously, a wavefront in a homogeneous and isotropic medium arrives first at the nearest accelerometer. Thus, the wavefront arrives first at A_1 when touching the point "b". The wave arrives first at A_2 touching point "d". The wavefront reaches simultaneously at the two accelerometers pressing the middle point "c".

2.3. Delay estimation via NCC

We use normalized cross correlation (NCC) to compute the relative delay between the signals acquired by the two accelerometers (see Appendix A for the definition and computation of NCC). Suppose that the wavefront generated by a keystroke takes n_1 sampling periods to reach the accelerometer A_1 and takes n_2 sampling periods to reach the accelerometer A_2 (see Fig. 2). In this case, we will observe a peak in NCC between the acceleration values obtained by A_1 and those obtained by A_2, when the latter is shifted right $n_1 - n_2$ positions. This difference is the estimated delay.

The right column of Fig. 2 depicts the NCC between the signals acquired by the two accelerometers touching the acrylic plate with the finger. Using the peaks in NCC we computed the group velocity, that was estimated as ≈45m/s. Fig. 3 depicts the signals obtained and the NCC touching the plate with the tip of a mechanical pencil. The group velocity is more than twice faster, ≈95m/s, because the frequency generated touching the plate with the pencil is higher than touching it with the finger.

The duration of the first semi-cycle of the signal $A1_z$ tapping with the finger (Fig. 2) is ≈5ms, corresponding to frequency of ≈100Hz (if considered cyclic). The duration of the first semi-cycle of the signal $A1_z$ tapping with the pencil (Fig. 3) is ≈1ms, corresponding to frequency of ≈500Hz (if considered cyclic).

3. Acrylic Plate Mockup Keypad

We constructed a mockup keypad using an acrylic plate to verify if the vibration delay can be used to identify the pressed key. We fixed a paper print of a keypad on the plate (Fig. 4), glued three accelerometers and touched inside each region emulating the keys. If we achieve a high accuracy in this test, it would be worth continuing the tests in real devices. We pressed 10 times each one of "0" to "9" virtual keys, generating 100 acquisitions.

Fig. 5 (top) depicts a typical keystroke captured by the three accelerometers. These signals are complex due to dispersion, reflections and many other wave

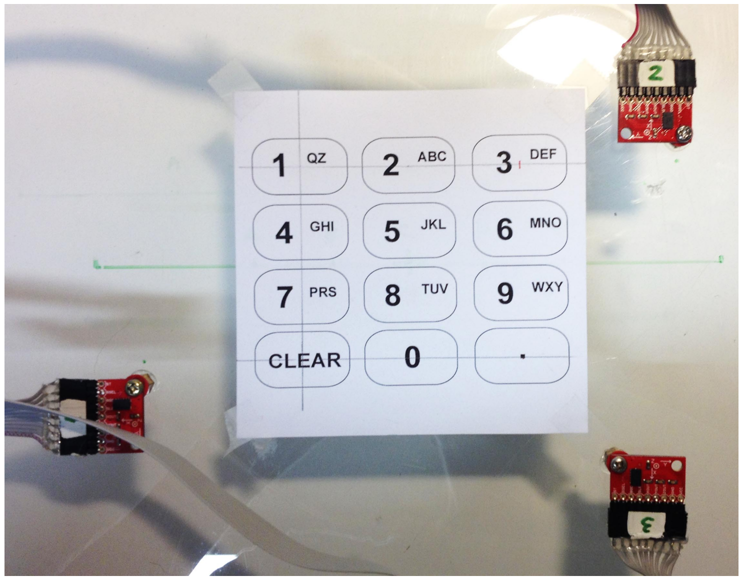

Figure 4: Keypad emulation on an acrylic plate. We tapped inside each virtual key to emulate keystrokes.

phenomena. So, if we simply compute the NCC between a pair of these signals, the highest peak may not correspond to the relative delay. It is possible to remove the artifacts introduced by the reflections by analyzing only the first points in time of the signal, before the arrival of the reflections. To this end, we enveloped the signals with a Gaussian window with mean μ and standard-deviation σ. We compute the highest peak in the first M points in each of the three original signals and then set μ as the average position of the three peaks, as shown in Fig. 5. The parameters σ and M depend on the experiment.

In this experiment, the sampling rate was 25KS/s or 50KS/s. We used $\sigma = \frac{200}{\sqrt{2}}$ for sampling rate of 25KS/s and $\sigma = \frac{100}{\sqrt{2}}$ for 50KS/s, and $M = 1500$ for both.

After multiplying the three original signals with the Gaussian window, we take pairs of the enveloped signals and compute NCC between each pair. The position of the highest peak in NCC is considered the relative delay between the pair. As we have three original signals, we get three relative delays. However, we

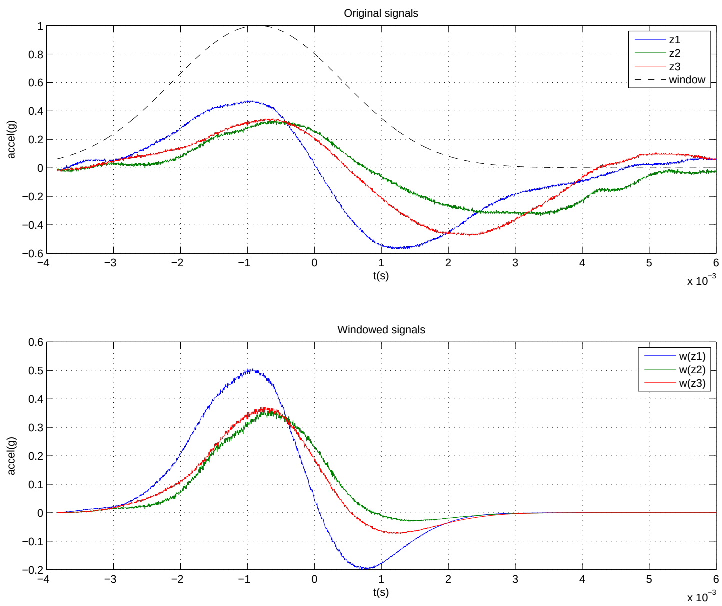

Figure 5: We apply a Gaussian window to attenuate the reflections and improve the delay estimation. (Top) the original acceleration signals took from the PIN-pad experiment, key "8". (Bottom) the windowed signals.

noted that only two out of these three relative delays are independent features, because the third can be obtained as a linear combination of the first two. See Appendix B for explanation.

In our previous paper [2], we extracted up to 165 features from each keystroke, instead of only two. The features were the values of NCC (instead of the position of the highest peak in NCC). In our very preliminary conference paper [19], we used many tentative features before choosing NCC. At those times, we made these choices because we had no clear idea of the underlying physical phenomenon.

We use in all experiments a simple Naive Bayes classifier with normal distribution. We took randomly 80% of all features as the training set and 20% as the test set, repeat this procedure 100 times and present the classification result as a confusion matrix.

In this "mockup keypad" experiment, we obtained 100% of correct classifi-

Table 1: Confusion Matrix of the Acrylic Plate Mockup Keypad Experiment

Key	1	2	3	4	5	6	7	8	9	0
1	200									
2		200								
3			200							
4				200						
5					200					
6						200				
7							200			
8								200		
9									200	
0										200

cation rate (Table 1)! Evidently, this is an ideal situation. In order to visualize spatially the data, we used 2-D trilateration (Appendix B) to estimate the relative locations of keys (Fig. 6). The estimated positions closely resemble their actual positions. Moreover, the clusters of keys are clearly separated. This shows that the group velocity is constant throughout the acrylic plate, because it is a homogeneous and isotropic medium as we assumed in the source localization method.

4. PIN-pad

After the experiment with the acrylic plate, we applied the vibration delay attack to a PIN-pad designed to deal with sensitive information in a secure way. Fig. 7 shows the device, an Ingenico iPP320 PIN-pad and the assembly of the experiment, where the three accelerometers were glued inside the SAM (Secure Access Module) card access compartment. This device is PCI-PTS compliant[1], under 2.X and 3.X versions.

[1] PCI stands for Payment Card Industry. PTS stands for PIN Transaction Security, a set of requirements specific for PIN entry devices. Device compliance can

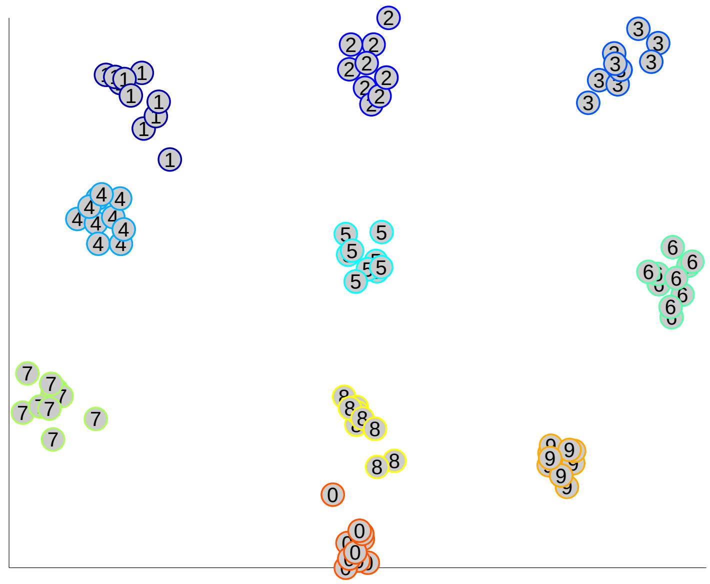

Figure 6: Estimation of the key positions of the acrylic keypad experiment.

The choice of this model/brand was not guided by any prior vulnerability we could spot. We want to make it clear that most of PIN-pads with a SAM card compartment are potential targets of this attack and Ingenico's iPP320 is not a special case.

The SAM card compartment increases the vulnerability to vibration delay attack mainly because:

1. it provides room for implanting wiretap devices or "bugs", hidden within the compartment;

2. the compartment is normally located just below the keypad, the ideal place to capture the vibrations from the keystrokes;

3. the SAM slots can eventually provide electrical power for the "bugs".

be consulted at `https://www.pcisecuritystandards.org/approved_companies_providers/approved_pin_transaction_security.php`

(a) (b)

Figure 7: (a) PIN-pad used in the experiment with approximate locations of accelerometers. (b) The bottom view showing the SAM compartment with the implanted accelerometers.

Thus, the attack can be executed in real scenario in a noninvasive and undetectable way, without batteries and wires. We placed the accelerometers inside the SAM card compartment to simulate a real vibration delay attack. In a real attack, however, miniaturized bug devices may be placed inside this compartment.

The restricted space in the SAM card compartment does not allow us to place the accelerometers wherever we want. So, the triangle formed by the three accelerometers covered only a small portion of the area where the keys are located (Fig. 7 (a)). We used spacers between the device's chassis and the printed circuit boards of the accelerometers, to make the accelerometers feel the vibration of only a small area, hoping that this may improve the results.

We pressed 20 times each one of the "0" to "9" keys. As before, we enveloped the three signals with a Gaussian window with $\sigma = \frac{450}{\sqrt{2}}$ and $M = 1500$. We used sampling rate of 250KS/s.

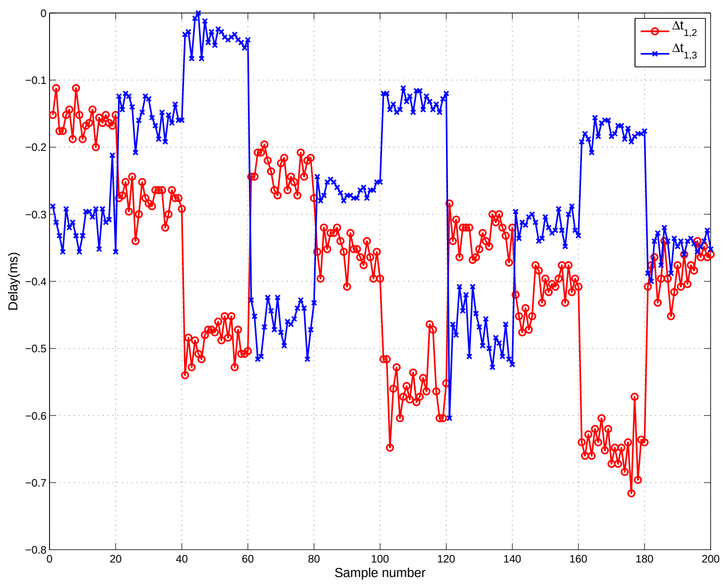

Figure 8: The features obtained in PIN-pad experiment. Each graph represents the delays between pairs of signals. X-axis is the test number, that is, $x = 1$ to 20 correspond to key "1", $x = 21$ to 40 to key "2" and $x = 181$ to 200 to key "0". Only $\Delta t_{1,2}$ and $\Delta t_{1,3}$ were used. See text.

4.1. Features

Fig. 8 depicts the obtained features. Each graph represents the delays between the signals obtained by a pair of accelerometers. The x coordinate indicates the test number. For example, $x = 1$ to 20 correspond to the 20 strokes of key "1", $x = 21$ to 40 to key "2" and $x = 181$ to 200 to key "0". As before, we used only two features, $\Delta t_{1,2}$ and $\Delta t_{1,3}$.

Table 2 shows the confusion matrix. The recognition rate is very high (96.4±6%), where 6 is the standard deviation of the 100 cross validations. The errors occur only between the neighboring keys. Moreover, excluding the key "0" (that seems to be a special case) the errors occur only between neighboring keys that belong to the same column. We observed similar results in [2]. Our hypothesis is that this happens because the distance between columns (≈23mm) is almost twice the distance between rows (≈13mm), making it easier to make

Table 2: Confusion Matrix of the Pin-pad Experiment

Key	1	2	3	4	5	6	7	8	9	0	Acc.(%)
1	400										100.0
2		398			2						99.5
3			400								100.0
4				400							100.0
5					397			3			99.2
6						382			18		95.5
7							389			11	97.2
8								345		55	86.2
9									400		100.0
0							5	50		345	86.2

row misclassifications.

In this experiment, the reconstruction of the key locations is also very good (Fig. 9) though the clusters are not so clearly separated as in the ideal case (Section 3). The keys are uniformly distributed in space with the exception of keys “0” and “8” that are partially mixed (in agreement with the confusion matrix in Table 2 and features in Fig. 8). These results show that the supposition of constant group velocity used in the source location method (Appendix B) is reasonable in practice, in spite of the apparent complexity of the medium. The observed localization errors may be due to: (a) the triangle formed by the accelerometers covers only a small part of the keypad; (b) the group velocity may not be constant throughout all the device; (c) the delay estimation method is not accurate enough; and (d) the medium is dispersive.

4.2. From NCC to TDOA

The instant of the peak in NCC can be used to estimate the delay between two similar signals (Section 2.3). In a previous work [2], we used the amplitudes of NCC as features to identify the pressed keys, without computing the instant of the peak. This implied large feature vectors (up to 165 features), as opposed to small TDOA features here used (only 2 features). However, the dimension

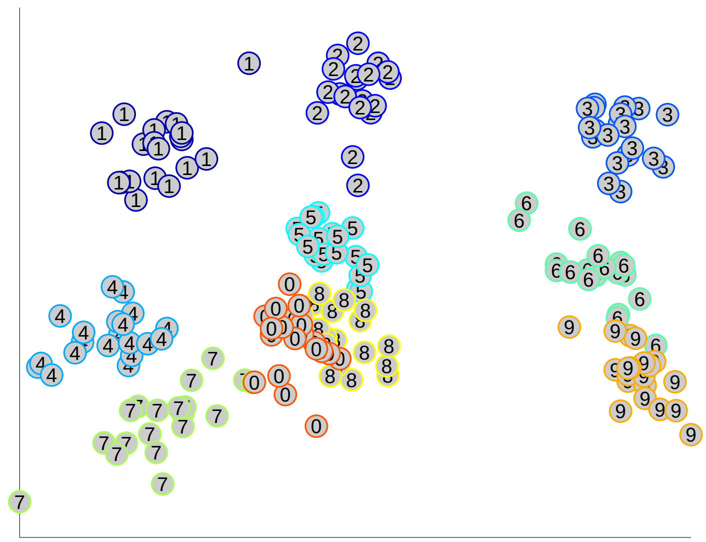

Figure 9: Estimation of the key locations of the PIN-pad experiment.

of NCC features can be reduced using some standard dimensionality reduction technique, as PCA (Principal Component Analysis). We used PCA to reduce NCC amplitudes into two main features. Fig. 10 shows that the two features so obtained are very well-correlated with the TDOA features we used throughout this paper. This clearly demonstrates the delay of arrival is the main physical phenomenon that identifies the pressed key.

Feeding the Bayes learning algorithm with the two features obtained by NCC-PCA, the obtained recognition rate was 95.1%, very close to the 96.4% obtained with the TDOA features. Using the three most important features obtained by NCC-PCA, the recognition rate is 97.7%, slightly higher than the rate obtained with the two TDOA features. This may indicate that there are other information (besides the time of arrival) in the NCC amplitudes that may help increasing slightly the recognition rate. Maybe the classifier is using specific vibration pattern of each key, wave reflections inside the device or some other complex phenomena to improve the key classification rate. If we use four features, the recognition rate decreases to 96.5%.

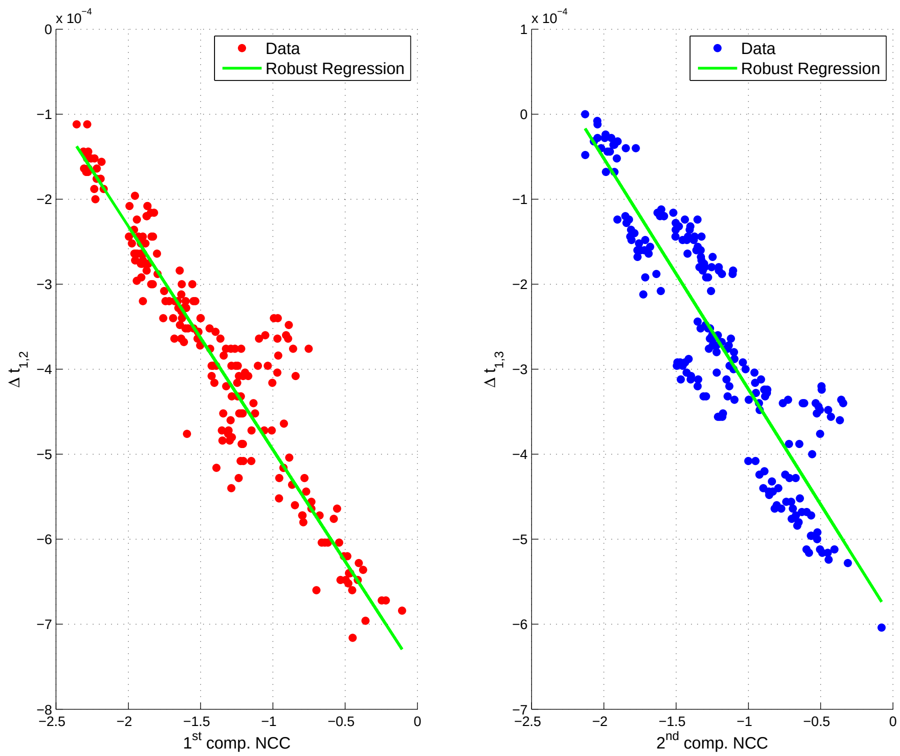

Figure 10: Comparison between the two main NCC-PCA components and the two TDOA features (the delays between the vibration signals). The pairs of features computed in different ways are highly correlated.

4.3. PCI Requirements

PCI requires that an attack such as described in this work should be possible only with very high cost of 26 for identification and 13 for exploitation[2]. Nevertheless, the vibration delay attack to this PCI-PTS compliant equipment costs only 12.5 for identification and 3.5 for exploitation (Table 3). The method

---

[2] "There is no feasible way to determine any entered and internally transmitted PIN digit by monitoring sound, electro-magnetic emissions, power consumption or any other external characteristic available for monitoring – even with the cooperation of the device operator or sales clerk – without requiring an attack potential of at least 26 for identification and initial exploitation with a minimum of 13 for exploitation." [20, p. 16]

Table 3: Calculation of the Cost of PIN-pad Vibration Delay Attack

Factor	Identification	Exploitation
Attack time	Beyond 160 hours = 5.5	$\leq$1 hour = 0
Expertise	Expert = 4	Layman = 0
Knowledge of the PIN entry device	Public = 0	Public = 0
Access to the PIN entry device	Mechanical Sample = 1	Mechanical Sample = 1
Equipment required for the attack	Standard = 1	Standard = 1
Specific parts required	Standard = 1	Standard = 1
Total cost	**12.5**	**3**

Table 4: Comparisons between the Main Experiments

	Previous work		This work
	ATM	**PIN2 rigid**	**PIN-pad**
Features per keystroke	63	165	3
Accelerometers	3	2	3
Success rate	98.4%	76.7%	96.4%

used to calculate the costs can be found in [20, p. 142].

5. Considerations

Table 4 compares the main experiments of our previous work [2] and of this work. Clearly, "PIN2 rigid mode" experiment of the previous work has the lowest success rate. As we now know the main physical phenomenon for leaking the information, we can explain the cause for this low rate. It is because in that experiment we used only two accelerometers. Thus, TDOA cannot uniquely determine the location of the vibration source.

The success rate of the ATM experiment of our previous work is slightly higher than PIN-pad experiment of this work. The recognition rates of the two experiments cannot be compared directly, because they are attacking two different devices (ATM keypad and PIN-pad). However, as we said in the last Section, it seems that there are other information (besides TDOA) in NCC amplitudes that may increase slightly the recognition rate.

6. Conclusion

In this paper, we have demonstrated that the primary cause that makes it possible to identify the pressed key by monitoring the vibrations with accelerometers is the relative delays in the wavefront arrival times at different accelerometers located at different points. We have shown that the propagation delay of the wavefront generated by the keystroke makes each accelerometer feel similar vibrations at different moments. These relative delays is used in our "vibration delay attack". A simple classification scheme using the relative delays yielded 96.4% of key recognition success rate.

We also have shown that a PIN-pad, a device properly designed to counter side-channel attacks and PCI-PTS compliant is very vulnerable to the vibration delay attack. Clearly, the vibration delay attack can also be applied to touch screen devices.

Our finding indicates (i) the care that an engineer must have to design secure human-machine interface devices in the future and (ii) a new attack vector that certification processes must address hereafter.

Appendix A. Normalized cross correlation

Let the vector v with elements v_i, $0 \leq i < N$ represent the acceleration values captured by an accelerometer. The mean-corrected vector $\tilde{v}$ has elements $\tilde{v}_i = v_i - \bar{v}$, where $\bar{v}$ is the mean of v. We use only mean-corrected acceleration values, because we are not interested in the static acceleration of gravity. The correlation coefficient between the two mean-corrected vectors is:

$$corr(\tilde{v}, \tilde{w}) = \frac{\tilde{v} \cdot \tilde{w}}{\|\tilde{v}\|\|\tilde{w}\|} \tag{A.1}$$

Correlation coefficient measures the "similarity" between the two vectors, invariant to bias (because the vectors are mean-corrected) and to gain (because the vectors are divided by their norms).

Normalized cross correlation (NCC) between vectors v and w is a vector denoted as $\tilde{v} \otimes \tilde{w}$ whose elements are the correlation coefficients computed between time-shifted vectors, ignoring the elements that do not have the matching pair. It has $2N-1$ elements:

$$(\tilde{\mathbf{v}} \otimes \tilde{\mathbf{w}})_n = \begin{cases} \dfrac{\displaystyle\sum_{i=0}^{N-n-1} \tilde{v}_i \tilde{w}_{n+i}}{\sqrt{\displaystyle\sum_{i=0}^{N-n-1} \tilde{v}_i^2 \sum_{i=0}^{N-n-1} \tilde{w}_{n+i}^2}}, & 0 \leq n < N \quad \text{(A.2a)} \\ (\tilde{\mathbf{w}} \otimes \tilde{\mathbf{v}})_{-n}, & -N < n < 0 \quad \text{(A.2b)} \end{cases}$$

Note that $(\tilde{v} \otimes \tilde{w})_0 = corr(\tilde{v}, \tilde{w})$. NCC has been used for a long time in computer vision to find templates in search images, in an operation called template matching. We use Matlab function `xcov(v,w,'coeff')` to compute NCC.

Appendix B. Source location estimation

We present here the technique used to estimate the positions of keys throughout this paper. Note that it is not necessary to know the spatial position of the pressed key in order to identify it. We suppose that the group velocity is constant throughout the device and consequently that the relative distances are roughly equivalent to the measured relative delays. For instance, we consider that the distance $d_1 - d_2$ is approximately equal to the measured relative delay $\Delta t_{1,2}$ between accelerometers A_1 and A_2 (Fig. B.11). Similarly, we assume that $d_1 - d_3 \approx \Delta t_{1,3}$ and $d_2 - d_3 \approx \Delta t_{2,3}$.

The following reasoning demonstrates that only two time differences carry useful information. Consider $\Delta t_{1,2} = t_1 - t_2$. Doing the same for $\Delta t_{1,3}$ and $\Delta t_{2,3}$, it is easy to see that $\Delta t_{2,3} = \Delta t_{1,3} - \Delta t_{1,2}$, a linear combination of the other two features, not carrying new information.

We estimate the source location P by a simple numerical optimization, using Matlab function `fminunc`. We minimize the following functional:

$$f = c_{1,2} + c_{1,3}. \tag{B.1}$$

where $c_{1,2}$ is:

$$c_{1,2} = \begin{cases} \left[d_1^2 - (d_2 + \Delta t_{1,2})^2\right]^2, & \Delta t_{1,2} \geq 0 \quad \text{(B.2a)} \\ \left[d_2^2 - (d_1 + \Delta t_{1,2})^2\right]^2, & \Delta t_{1,2} < 0. \quad \text{(B.2b)} \end{cases}$$

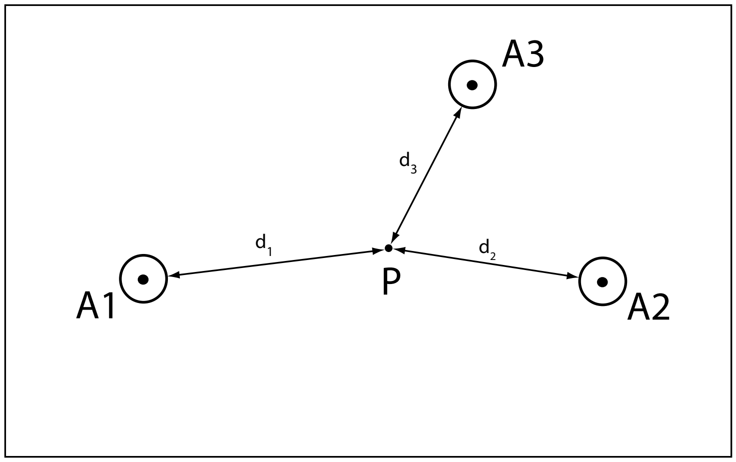

Figure B.11: Diagram of trilateration method.

The cost $c_{1,3}$ is defined similarly. The distance d_i from the accelerometer $A_i = (A_i^x, A_i^y)$ to the point $P = (P^x, P^y)$ is:

$$d_i = \|A_i - P\| = \sqrt{(A_i^x - P^x)^2 + (A_i^y - P^y)^2} \tag{B.3}$$

If we substitute Eq. B.2 and B.3 in Eq. B.1 and minimize f, we get the approximate position of point P (Fig. B.11).

References

[1] Verizon, Verizon 2014 Data Breach Investigations Report (2014).
URL `http://www.verizonenterprise.com/DBIR/2014/`

[2] G. de Souza Faria, H. Y. Kim, Identification of pressed keys from mechanical vibrations, Information Forensics and Security, IEEE Transactions on 8 (7) (2013) 1221–1229. `doi:10.1109/TIFS.2013.2266775`.

[3] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, 2005. `arXiv:http://epubs.siam.org/doi/pdf/10.1137/1.9780898717921`, `doi:10.1137/1.9780898717921`.
URL `http://epubs.siam.org/doi/abs/10.1137/1.9780898717921`

[4] D. Asonov, R. Agrawal, Keyboard acoustic emanations, in: Security and Privacy, 2004. Proceedings. 2004 IEEE Symposium on, 2004, pp. 3–11. doi:10.1109/SECPRI.2004.1301311.

[5] Y. Berger, A. Wool, A. Yeredor, Dictionary attacks using keyboard acoustic emanations, in: In Proceedings of Computer and Communications Security (CCS, 2006.

[6] L. Zhuang, F. Zhou, J. D. Tygar, Keyboard acoustic emanations revisited, ACM Trans. Inf. Syst. Secur. 13 (1) (2009) 3:1–3:26. doi:10.1145/1609956.1609959.
URL http://doi.acm.org/10.1145/1609956.1609959

[7] T. Halevi, N. Saxena, A closer look at keyboard acoustic emanations: Random passwords, typing styles and decoding techniques, in: Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security, ASIACCS '12, ACM, New York, NY, USA, 2012, pp. 89–90. doi:10.1145/2414456.2414509.
URL http://doi.acm.org/10.1145/2414456.2414509

[8] P. Marquardt, A. Verma, H. Carter, P. Traynor, (sp)iphone: Decoding vibrations from nearby keyboards using mobile phone accelerometers, in: Proceedings of the 18th ACM Conference on Computer and Communications Security, CCS '11, ACM, New York, NY, USA, 2011, pp. 551–562. doi:10.1145/2046707.2046771.
URL http://doi.acm.org/10.1145/2046707.2046771

[9] M. Ge, Analysis of source location algorithms - part i: Overview and non-iterative methods, J. Acoustic Emission 21.
URL http://www.ndt.net/article/jae/papers/21-014.pdf

[10] M. Ge, Analysis of source location algorithms - part ii: Iterative methods, J. Acoustic Emission 21.
URL http://www.ndt.net/article/jae/papers/21-029.pdf

[11] K. Ho, Y. Chan, Solution and performance analysis of geolocation by tdoa, Aerospace and Electronic Systems, IEEE Transactions on 29 (4) (1993) 1311–1322. doi:10.1109/7.259534.

[12] F. Gustafsson, F. Gunnarsson, Positioning using time-difference of arrival measurements, in: Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on, Vol. 6, 2003, pp. VI–553–6 vol.6. doi:10.1109/ICASSP.2003.1201741.

[13] D. Manolakis, Efficient solution and performance analysis of 3-d position estimation by trilateration, Aerospace and Electronic Systems, IEEE Transactions on 32 (4) (1996) 1239–1248. doi:10.1109/7.543845.

[14] T. Schumacher, D. Straub, C. Higgins, Toward a probabilistic acoustic emission source location algorithm: A bayesian approach, Journal of Sound and Vibration 331 (19) (2012) 4233 – 4245. doi:http://dx.doi.org/10.1016/j.jsv.2012.04.028.
URL http://www.sciencedirect.com/science/article/pii/S0022460X12003446

[15] K. Arun, E. Ong, A. Khong, Source localization on solids using kullback-leibler discrimination information, in: Information, Communications and Signal Processing (ICICS) 2011 8th International Conference on, 2011, pp. 1–5. doi:10.1109/ICICS.2011.6174299.

[16] W. Elmore, M. A. Heald, Physics of Waves, Dover, 1969.

[17] L. Meirovitch, Analytical Methods in Vibrations, Macmillan, 1967.

[18] Freescale Semiconductor - MMA7361LC: ±1.5g, ±6g, 3-Axis Analog Output Acceleration Sensor (February 2015).
URL http://www.freescale.com/webapp/sps/site/prod_summary.jsp?code=MMA7361LC

[19] G. S. Faria, H. Y. Kim, Identificação das teclas digitadas a partir da vibração mecânica, in: Anais do 30º Simpósio Brasileiro de Telecomunicações - SBrT, 2012.

[20] Payment Card Industry - Security Standards Council LLC, PIN Transaction Security (PTS) Point of Interaction (POI) Modular Derived Test Requirements v4.0 (June 2013).