Proximity Assurances Based on Natural and Artificial Ambient Environments

Iakovos Gurulian[1(✉)], Konstantinos Markantonakis[1], Carlton Shepherd[1], Eibe Frank[2], and Raja Naeem Akram[1]

[1] Information Security Group Smart Card Centre, Royal Holloway, University of London, Egham, UK
{Iakovos.Gurulian.2014,k.markantonakis,Carlton.Shepherd.2014, r.n.akram}@rhul.ac.uk

[2] Department of Computer Science, University of Waikato, Hamilton, New Zealand
eibe@waikato.ac.nz

Abstract. Relay attacks are passive man-in-the-middle attacks that aim to extend the physical distance of devices involved in a transaction beyond their operating environment. In the field of smart cards, distance bounding protocols have been proposed in order to counter relay attacks. For smartphones, meanwhile, the natural ambient environment surrounding the devices has been proposed as a potential Proximity and Relay-Attack Detection (PRAD) mechanism. These proposals, however, are not compliant with industry-imposed constraints that stipulate maximum transaction completion times, e.g. 500 ms for EMV contactless transactions. We evaluated the effectiveness of 17 ambient sensors that are widely-available in modern smartphones as a PRAD method for time-restricted contactless transactions. In our work, both similarity- and machine learning-based analyses demonstrated limited effectiveness of natural ambient sensing as a PRAD mechanism under the operating requirements for proximity and transaction duration specified by EMV and ITSO. To address this, we propose the generation of an Artificial Ambient Environment (AAE) as a robust alternative for an effective PRAD. The use of infrared light as a potential PRAD mechanism is evaluated, and our results indicate a high success rate while remaining compliant with industry requirements.

Keywords: Mobile payments · Relay attacks · Ambient environment sensing · Contactless · Experimental analysis

1 Introduction

Today, a wide variety of application environments exist that demand proximity of a user with a physical terminal, as well as high throughput, i.e. maximising the number of transactions per unit time. Both smart card-based payments and transport-related transactions are major examples of such applications in everyday life. These particular services are governed by industry-accepted standards,

P. Farshim and E. Simion (Eds.): SecITC 2017, LNCS 10543, pp. 83–103, 2017.
https://doi.org/10.1007/978-3-319-69284-5_7

such as the EMV specifications for card and mobile contactless payments. Under EMV, contactless transactions should complete within 500 ms [2–4]. Similarly, transport-related transactions should complete between 300 and 500 ms [1]. In addition to these, other applications exist that depend on proximity and transaction time, particularly in the realm of the Internet of Things (IoT), such as taking medical equipment inventories in operating theatres. The domain of sensor networks is another closely-related area where communication time and the proximity of sensors can be of paramount importance.

In this paper, we examine the problem of proximity detection in applications with restricted time-frames. Specifically, we focus on applications that are deployed traditionally as contactless smart cards but are gradually migrating to mobile phones using Near-Field Communication (NFC). During an NFC-based mobile contactless transaction, a mobile handset is brought into the radio range (<3 cm) of a payment terminal through which a dialogue is initiated. NFC, however, has no provisions to ascertain whether the device is genuinely in proximity to the terminal, which makes them susceptible to relay attacks.

In a relay attack [8,9,38], the aim of the malicious actor is to extend the physical distance of the communication channel between the victim's mobile phone and the transaction terminal – relaying each message across this extended distance. The attacker extends this distance using equipment that masquerades as legitimate devices to both the terminal and victim device, as shown in Fig. 1. The attacker has the potential to gain access to services using the victim's account if messages are relayed successfully without detection. At present, additional security mechanisms, like fingerprint scanning and Personal Identity Number (PIN) entry, may also be required in order to perform a contactless mobile transaction for a payment, transport ticketing, and similar services. However, even the use of PINs and biometrics cannot always prevent relay attacks (see the Mafia fraud attack [7]).

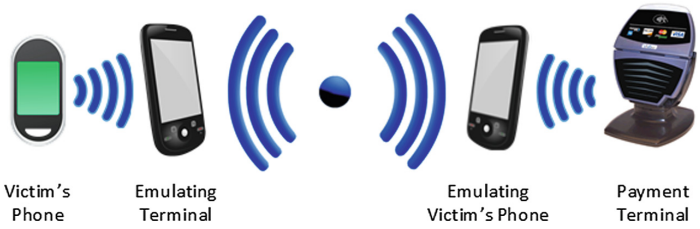

Fig. 1. Overview of a relay attack [31].

In recent years, a deluge of Proximity/Relay Attack Detection (PRAD) mechanisms have been proposed that rely on collecting information regarding the ambient environment surrounding the transaction instrument and terminal. Such proposals collect data using the sensors in modern mobile devices – such as temperature, motion and position sensors – which is compared for similarity to assure

that the transaction devices were genuinely in proximity. In this work, we present an empirical evaluation of the claim that *ambient sensing on mobile devices is an effective PRAD method under the time conditions stipulated by industry*. We present an extended study before proposing the utility of an artificially generated ambient environment as an alternative PRAD mechanism.

2 Natural Ambient Environment Sensing

In this section, we discuss a number of generic deployment models for deploying proximity- and transaction time-sensitive applications using ambient sensing. Next, we discuss related work before evaluating the claim that ambient sensors are an effective PRAD mechanism under the real-world constraints imposed by industry requirements, i.e. by EMV and ITSO.

2.1 Ambient Sensors in Conventional Transactions

For contactless smart cards, relay attacks can be countered using distance bounding protocols [27] and variants of such [18]. This is still an active research domain, with new attacks and countermeasures emerging [5,7,17]. At the current state of the art, however, these are not easily transferable to NFC-enabled phones, due to their high sensitivity to time delays [6,16,35]. Alternative methods have been proposed to provide proximity detection, most of which use environmental and motion sensors present on modern mobile handsets [16,23,32,34,36,37]. In Sect. 2.2, we discuss how ambient sensors have been proposed to counter relay attacks in NFC-based mobile contactless transactions.

An ambient sensor measures a particular environmental property of its immediate surroundings, such as temperature, light, humidity and sound; a wealth of such sensors are deployed in modern smartphones and tablets. In Fig. 2, we illustrate a generic approach for deploying ambient sensing as a proximity detection mechanism for mobile payments, with the following variations:

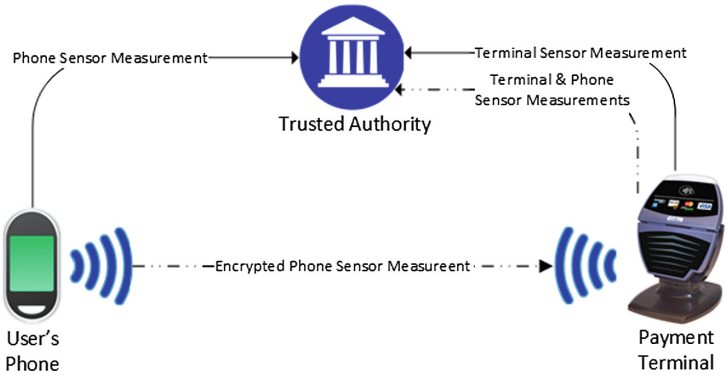

Fig. 2. Generic deployment of mobile sensing for proximity detection [31].

1. **Independent Reporting.** Both the smartphone and payment terminal collect sensor measurements independently of each other and transmit these to a trusted authority (depicted as solid lines in Fig. 2). The authority compares the sensor measurements, based on some predefined comparison algorithm with set margins of error (threshold), and decides whether the two devices are within proximity to each another.
2. **Payment Terminal Dependent Reporting.** This set-up involves the smartphone encrypting the sensor measurements with a shared key between smartphone and trusted authority, and transmitting the encrypted message to the payment terminal (shown as double-dot-dash lines in Fig. 2). The payment terminal sends its own measurements and the smartphone's to the trusted authority for comparison.
3. **Payment Terminal (Localised) Evaluation.** The smartphone transmits its measurement to the payment terminal, which compares it with its own measurements locally; the payment terminal then decides whether the smartphone is in proximity.

2.2 Related Work

We identify and summarise key pieces of related work that have suggested using natural ambient sensing as a PRAD mechanism.

Ma et al. [23] explored the use of GPS (Global Positioning System) location data for determining the proximity of two mobile phones. A ten-second recording window was used in which GPS data was collected every second, which was subsequently compared across various devices. The work reported a high success rate in identifying co-located devices.

Halevi et al. [16] demonstrated the use of ambient sound and light for proximity detection. The authors analysed sensor measurements – collected for 2 and 30 s duration for light and audio respectively – using a range of similarity comparison metrics. Extensive experiments were performed in different physical locations with a high success rate in detecting proximate devices.

Varshavsky et al. [37] used the shared radio environment of devices – the presence of WiFi access points and associated signal strengths – as a proximity detection mechanism for secure device pairing. The approach was considered to produce low error rates and, while it did not focus on NFC-based mobile transactions, their techniques and methodologies may still be applicable.

Urien et al. [36] suggested using ambient temperature with an elliptic curve-based RFID/NFC authentication protocol to determine whether two devices are co-located before creating a secure channel. The proposal combines the timing channels in RFID, used traditionally in distance bounding protocols, in conjunction with ambient temperature. The work, however, was not implemented and has no experimental data to evaluate its effectiveness.

Mehrnezhad et al. [25] proposed the use of an accelerometer to provide proximity assurances of a mobile device with a payment terminal. The scheme requires the user to tap the terminal twice in succession, before comparing the sensor data from the device and the terminal for similarity. It is difficult to

deduce the total time it took to complete a transaction entirely, but the authors provide a recording time range of 0.6–1.5 s.

Truong et al. [34] evaluated four different sensors using a recording duration of 10–120 s. While the results were positive, the long sampling duration renders it unsuitable for NFC-based mobile transactions.

Additional work by Jin et al. [20] showed that a smartphone's magnetometer can be used to establish proximity assurance. This approach requires more than 500 ms; the authors do not claim that magnetometers can provide an effective relay attack detection mechanism.

Shrestha et al. [32] used a number of ambient sensors within specialised hardware, known as Sensordrone, for proximity detection. The work was not evaluated using the ambient sensors available on commodity handsets, did not provide the sampling duration, and states that data from each sensor was collected for a few seconds.

Table 1. Related work in sensor-based PRAD mechanisms.

Paper	Sensor(s) used	Sample duration	Contactless suitability
Ma et al. [23]	GPS	10 s	Unlikely
Halevi et al. [16]	Audio	30 s	Unlikely
	Light	2 s	More Likely
Varshavsky et al. [37]	WiFi (Radio waves)	1 s	More Likely
Urien et al. [36]	Temperature	N/A	–
Mehrnezhad et al. [25]	Accelerometer	0.6 to 1.5 s	More Likely
Truong et al. [34]	GPS raw data	120 s	Unlikely
	WiFi	30 s	Unlikely
	Ambient audio	10 s	Unlikely
	Bluetooth	12 s	Unlikely
Shrestha et al. [32]	Temperature (T)	Few seconds	Unlikely
	Precision Gas (G)	Few seconds	Unlikely
	Humidity (H)	Few seconds	Unlikely
	Altitude (A)	Few seconds	Unlikely
	HA	Few seconds	Unlikely
	HGA	Few seconds	Unlikely
	THGA	Few seconds	Unlikely

Table 1 summarises past work, using sensor sampling durations to determine their suitability for NFC-based mobile phone transactions in banking and transportation. 'Unlikely' proposals have sample durations so large that they may not be adequate for mobile-based services that substitute contactless cards, while those with reasonably short durations are labelled 'More Likely'. However, even

schemes denoted as 'More Likely' may not be suitable as no proposal is evaluated under the time constraints stipulated by the banking and transport sectors. In these sectors, the goal is to serve people as quickly as possible to maximise customer throughput, as alluded to in Sect. 1; time is critical in determining whether a transaction is successful and, indeed, permitted. Here, an optimal transaction duration is 500 ms rather than seconds.

Our initial study (Shepherd et al. [31]) questioned the effectiveness of ambient sensing as a proximity detection mechanism under short time frames (500 ms) – illustrating that numerous sensors available via the Android platform perform poorly within an operating distance of <3 cm and transaction-duration of <500 ms. Both threshold- and machine learning-based analyses were employed using sensor data collected from mock transactions in the field. Similar results were also exhibited by further experimentations (Haken et al. [15]) using sensors on the Apple iOS platform. Our third analysis (Gurulian et al. [14]) selected seven of the best-performing sensors from our first study [31]. In this study, sensor data from genuine and relay transactions was collected from an emulated relay attack set-up, with the goal of determining whether data from relayed transactions can be distinguished from legitimate ones. In the following sections, we reproduce the results from these initial studies along with additional analyses conducted post-publication. Note that the focus of this work is on conventional transactions that require no further interaction with the terminal, e.g. double-tapping, a gesture, or otherwise. Ambient sensing has also been used in various user-device authentication, key generation and secure channel schemes [21,30]. These applications typically measure the environment for longer periods of time (>1 s) and, generally speaking, their primary goal is not proximity detection of a device with a terminal. As such, we omit these from the discussion.

2.3 Approaches and Evaluation Metrics

In previous work, two approaches have been used predominately for sensing-based PRAD mechanisms:

- *Threshold-based Similarity*: the use of time and frequency domain similarity metrics, such as Mean Absolute Error (MAE), Pearson's Correlation Coefficient and Coherence. A single threshold is generated that aims to separate all legitimate transactions from illegitimate ones using a particular similarity metric. The transaction is accepted if the metric result falls within this pre-set threshold of the maximum allowed dissimilarity.
- *Machine Learning*: the use of well-known classification algorithms, such as Naïve Bayes, Support Vector Machines (SVMs) and Random Forests. The classifier is trained on a set of feature vectors with corresponding binary labels (legitimate or relayed transaction), which are collected beforehand. The trained model is used to classify subsequent transaction data streams as legitimate or relayed.

Standard binary classification evaluation metrics have been applied to measure the effectiveness of a particular scheme, namely classification accuracy [16],

f-scores[1] [32,34] and Equal Error Rate (EER) [25]. F-scores and EERs involve the computation of *false positives/acceptances* (the number of relayed transactions accepted erroneously) and *false negatives/rejections* (the number of legal transactions rejected). F-scores account for *precision*, the correct positive results divided by the number of all positive results, and *recall*, the number of correct positive results as proportion of the number of positive results that should have been identified (see Eq. 1). The EER – used extensively in biometrics, e.g. fingerprint recognition [24] – is found by calculating the False Acceptance Rate (FAR) and False Rejection Rate (FRR), shown in Eq. 2, over a range of thresholds and finding the rate at which $FAR = FRR$. Alternatively, some authors have opted to present the FAR and FRR results alone [37]. Finally, accuracy represents the correct identification of positive and negative transactions in the test set (Eq. 3), but does not clearly illustrate the number of false positives and negatives.

F-scores and accuracy have been used to primarily evaluate machine learning-based relay attack detection, e.g. [32,34], while EERs have been employed for threshold-based similarity approaches [25] to find an acceptance threshold that, broadly speaking, balances usability (false rejection rate) with security (false acceptance rate). We use the EER as a common evaluation metric for assessing the performance of machine learning and threshold-based approaches across a variety of similarity metrics.

$$F_{score} = \frac{2TP}{2TP + FP + FN} \tag{1}$$

$$FAR = \frac{FP}{FP + TN} \qquad FRR = \frac{FN}{FN + TP} \tag{2}$$

$$Accuracy = \frac{TP + TN}{P + N} \tag{3}$$

2.4 Effectiveness for Proximity Detection

In our first study [31], we evaluated the effectiveness of ambient sensors to determine whether two devices are in proximity to one another (irrespective of whether a relay attack is in action). A field trial was conducted in which sensor data from 1000 transactions per sensor was collected from 252 users at four different locations on a university campus. Two devices were used for the data collection: a transaction terminal (TT), and a transaction instrument (TI). Data was collected for 500 ms upon the initiation of the NFC-based transaction, and stored locally for later evaluation (Fig. 3).

We subjected this data to the two analyses discussed in Sect. 2.3 – threshold-based similarity and machine learning – to determine whether data from legitimately co-located devices can be distinguished from non-proximate pairs. The implementation of the test-bed, data analysis and collected data sets are made available at: https://github.com/AmbientSensorsEvaluation/Ambient-Sensors-Proximity-Evaluation.git. The source code for the additional

[1] Also known as the F1 score or F-measure.

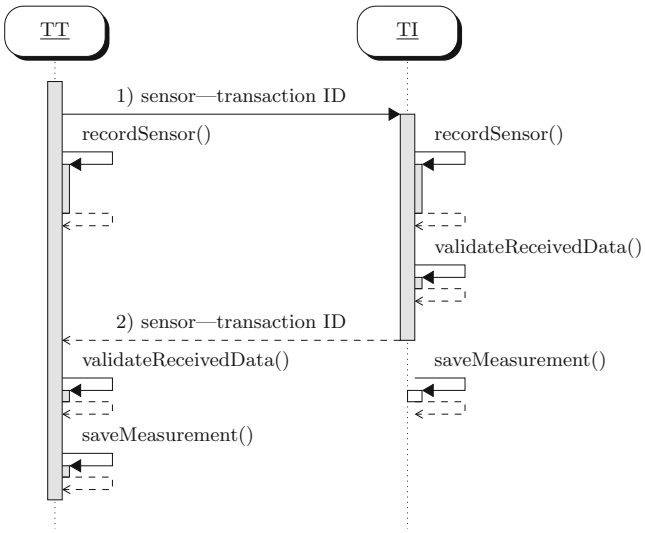

Fig. 3. Measurement recording overview.

threshold-based experiments presented in this paper can be found at: https://github.com/AmbientSensorsEvaluation/Threshold-Based-Analysis.

Analysis Approach. For each sensor, the EER and associated threshold, t, were computed using six time- and frequency-domain similarity measures, including those used in previous work. We list these forthwith. *Time domain metrics*: Mean Absolute Error (MAE), Eq. 4; Pearson's correlation coefficient [25], Eq. 5; maximum cross-correlation [16,34], Eq. 6; and Euclidean distance [34], Eq. 7. *Frequency domain*: coherence [25], Eq. 8. *Both domains*: time-frequency distance [16,34], Eq. 9. In [31], we presented results only from the MAE and Pearson's correlation coefficient; in this work, we present the results from all of these similarity metrics. Each metric was applied directly onto the sensor data collected during the field trials. For machine learning, the Weka package was employed, while a Python application was developed for threshold-based similarity learning using the Numpy, Scipy, Matplotlib and Pandas Python packages for metric implementations, graph plotting and CSV data processing.

$$MAE(A,B) = \frac{1}{N}\sum_{i=1}^{N}|A_i - B_i| \tag{4}$$

$$corr(A,B) = \frac{\sum_{i=1}^{N}((A_i - \mu_A)(B_i - \mu_B))}{\sqrt{\sum_{i=1}^{N}(A_i - \mu_A)^2 \sum_{i=1}^{N}(B_i - \mu_B)^2}} \tag{5}$$

Where μ_A represents the arithmetic mean of A.

$$M_{corr}(A,B) = \max(cross\_correlation(A,B)) \tag{6}$$

$$d(A,B) = \sqrt{\sum_{i=1}^{N} (B_i - A_i)^2} \tag{7}$$

$$C_{AB}(f) = \frac{|G_{AB}(f)|^2}{G_{AA}(f) \cdot G_{BB}(f)} \qquad F_{AB} = \sum_{f} C_{AB}(f) \tag{8}$$

Where G_{AA} is the auto-spectral density of A and G_{AB} is the cross-spectral density of signals A and B (left). The similarity is found by the sum of the magnitudes of coherence values at all frequencies (right).

$$Diff(A,B) = \sqrt{D_{time}(A,B)^2 + D_{freq}(A,B)^2} \tag{9}$$

Where $D_{time}(A,B) = 1 - M_{corr}(A,B)$ and $D_{freq}(A,B) = ||FFT(A) - FFT(B)||$, in which $||FFT(A) - FFT(B)||$ is the Euclidean norm of the FFTs of signals A and B.

Results. The results for the threshold-based and machine learning analyses are presented in Tables 2 and 3 respectively. Note that the proximity sensor was excluded from the analysis. On some Android devices, proximity sensors return the precise distance at which an object is located from the sensor, whereas others return a binary value for whether an object is close to/far from the sensor (within 5 cm)[2]. Our test-bed devices returned only binary values. Virtually every transaction contained 'far' values, as the devices were tapped back-to-back and the sensor was located on the front of the device. Consequently, this returned identical values in almost all cases when applying the similarity metrics described previously, e.g. $MAE = 0$, which impeded threshold-finding. Machine learning was able to capture the rare times in which the sensor returned 'close' values, like when the user covered the device with their hand during the transaction. While the machine learning results are included, the issues identified mean they should be treated with caution. Other technical challenges existed elsewhere; the Rotation Vector sensor, for example, returned significant numbers of zero values on the test-bed devices, which likely distorted the results of our analysis, while sound was capable of capturing values for only half of the permitted 500 ms time-frame. The reader is referred to [31] for a breakdown of sensor success and any technical limitations encountered.

The results indicate that no sensor in either analysis can satisfactorily distinguish between proximate and non-proximate device data pairs. Some sensors provide virtually no discrimination and perform similarly to a random classifier, e.g. accelerometer (43.4–49.8% EER) and linear acceleration (42.6–50.0%). Other sensors provide better discrimination, e.g. magnetic field (29.2–32.3%) and pressure (9.2–27.0%), but still fall short of acceptable performance. Even in the best case – the pressure sensor using the Decision Tree classifier – the

[2] http://developer.android.com/guide/topics/sensors/sensors_position.html#sensors-pos-prox.

Table 2. Threshold-based EERs for each sensor with Mean Absolute Error (MAE), Pearson's Correlation Coefficient (PCC), Maximum Cross-Correlation (C-Corr), Euclidean Distance (ED), Coherence (Coh) and Time-Frequency Distance (T-FD). Best result for each sensor shown in bold.

Sensor	MAE	PCC	C-Corr	ED	Coh	T-FD
Accelerometer	**0.434**	0.458	0.501	0.498	0.542	0.501
GRV[a]	**0.384**	0.486	0.500	0.442	0.524	0.498
Gravity	0.429	**0.424**	0.498	0.501	0.506	0.498
Gyroscope	0.443	**0.441**	0.493	0.498	0.548	0.499
Light	0.488	0.496	0.545	0.502	**0.471**	0.546
Linear acceleration	0.496	**0.426**	0.494	0.507	0.507	0.500
Magnetic field	**0.323**	0.384	0.537	0.337	0.568	0.536
Pressure	**0.270**	0.492	0.601	0.283	0.503	0.601
Rotation Vector	0.498	0.466	0.501	0.278	0.500	**0.273**
Sound	0.417	0.488	0.481	**0.338**	0.518	0.481

Proximity excluded due to insufficient unique values.
[a]GRV: Geomagnetic Rotation Vector sensor

Table 3. Estimated EERs for machine learning algorithms, obtained by repeating stratified 10-fold cross-validation 10 times. Best result for each sensor shown in bold.

Sensor	Random Forest	Naive Bayes	Logistic Regression	Decision Tree	Support Vector Machine	Multilayer Perceptron
Accelerometer	0.626 ± 0.024	0.509±0.026	0.526±0.023	0.500 ± 0.0	**0.498** ± 0.025	0.551± 0.025
GRV	**0.435**± 0.021	0.447 ± 0.024	0.474 ± 0.031	0.500 ± 0.0	0.489 ± 0.036	0.450 ± 0.026
Gravity	0.874 ± 0.018	0.579 ± 0.020	0.579 ± 0.024	**0.500** ± 0.0	**0.500** ± 0.026	0.746 ± 0.112
Gyroscope	0.683 ± 0.027	**0.499** ± 0.024	0.543 ± 0.024	0.500 ± 0.0	0.511 ± 0.025	0.514 ± 0.025
Light	0.576 ± 0.026	0.515 ± 0.024	0.533 ± 0.025	**0.500** ± 0.0	0.508 ± 0.024	0.513 ± 0.028
Linear acceleration	0.603 ± 0.025	0.507 ± 0.027	0.543 ± 0.023	**0.500** ± 0.0	**0.500** ± 0.021	0.554 ± 0.028
Magnetic field	**0.292** ± 0.021	0.319 ± 0.020	0.322 ± 0.020	0.415 ± 0.015	0.398 ± 0.046	0.329 ± 0.026
Pressure	0.103 ± 0.010	0.107 ± 0.010	0.287 ± 0.013	**0.092** ± 0.054	0.319 ± 0.045	0.114 ± 0.019
Proximity	0.499 ± 0.031	0.537 ± 0.069	**0.476** ± 0.188	0.500 ± 0.0	0.543 ± 0.254	0.508 ± 0.197
Rotation Vector	**0.276** ± 0.046	0.563 ± 0.243	0.596 ± 0.233	0.500 ± 0.0	0.513 ± 0.243	0.488 ± 0.245
Sound	**0.288** ± 0.019	0.314 ± 0.022	0.310 ± 0.021	0.347 ± 0.136	0.411 ± 0.041	0.306 ± 0.020

EER was 9.2%. By definition of the EER, this implies that approximately 9.2% of both legitimate and illegitimate transactions would be rejected and accepted respectively. Rejecting almost 1-in-10 legitimate transactions in a high throughput scenario is likely to cause user annoyance in practice, such as mobile ticketing in a subway system. As such, it is difficult to recommend any single sensor in our analysis as an effective proximity detection method.

2.5 Effectiveness for Relay Attack Detection

The first evaluation [31] focused only on proximity detection, rather than using data from relay attacks. In our next major work, we conducted further field trials [14] in which data was collected from two devices that were genuinely in proximity, and a third device that was located 1.5 m/5 ft away. This replicated a relay attack in which an adversary launches the attack on a nearby victim, such as in a shop queue. We aimed to determine whether sensor data from the relay device pair – the terminal and the device 5 ft away – could be distinguished from a legitimate pair, i.e. the terminal and the device in proximity.

The relay pair comprised a transaction terminal (TT) and a transaction instrument (TI′), whereas the legitimate pair consisted of a relay transaction terminal (TT′), and a transaction instrument (TI). Devices TI′ and TI were tapped simultaneously against devices TT and TT′ respectively. A 500 ms NFC-based transaction was then initiated on both sides and, upon completion, the devices TT, TI′, and TI stored the collected sensor data locally for off-line analysis. Figure 4 presents an overview of the recording process.

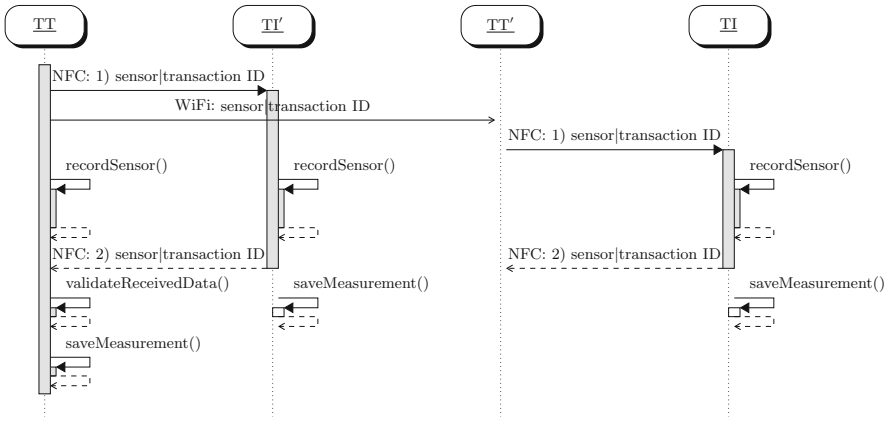

Fig. 4. Measurement recording process.

The implementation of the test-bed, data analysis and collection data sets for [14] are available at: https://github.com/AmbientSensorsEvaluation/Ambient-Sensors-Relay-Attack-Evaluation.

Analysis Approach and Results. In this study, we limited our sensor selection to the best performing sensors from our first analysis (Shepherd et al. [31]), i.e. those which successfully and consistently captured values within the 500 ms time limit over 1,000 transactions. Additionally, based on our initial investigations, we eliminated sensors that are largely uncommon on commodity handsets in the current market (2016–2017), like the pressure sensor. The reader is

referred to [14,31] for a detailed discussion of these matters. The same analysis techniques were used as the previous work for proximity detection – threshold-based analysis with the same similarity measures, and machine learning – using the EER evaluation metric. The results for the threshold-based and machine learning analyses of this study can be found in Tables 4 and 5 respectively.

Similar to our previous study (Sect. 2.4), some sensors provided poor discriminatory power; the magnetic field sensor, for instance, gave EERs of 36.1% and 43.3% in the analyses. Some sensors provided greater discriminatory power, e.g. gyroscope (17.9% EER with Random Forest) and the rotation vector sensor (27.7%, also with Random Forest). These EERs, however, are still too high to recommend as an effective PRAD in high throughput situations. Based on this evaluation, we reached the tentative conclusion that sensing the transaction devices' natural ambient environment may not be a suitable PRAD mechanism under industry-specified time constraints. In future work, we aim to conduct additional experiments using multiple sensors with various permutations and sensor fusion techniques to further interrogate the veracity of our conclusions.

Table 4. Threshold-based EERs (using the metric abbreviations in Table 2).

Sensor	MAE	PCC	C-Corr	ED	Coh	T-FD
Accelerometer	0.494	0.477	0.590	**0.468**	0.507	0.590
Gyroscope	0.521	**0.455**	0.535	0.495	0.528	0.489
Magnetic field	0.444	0.473	0.470	**0.433**	0.487	0.470
Rotation vector	0.330	0.472	**0.327**	0.670	0.534	0.509
Gravity	0.521	0.490	0.401	**0.289**	0.503	0.362
Light	**0.367**	0.488	0.444	0.372	0.505	0.437
Linear acceleration	0.482	0.536	0.503	0.506	**0.443**	0.493

The poor performance of measuring the natural environment as a PRAD led us to explore the generation of an artificial ambient environment that is unique to each transaction. In the following section of this paper, we introduce and discuss artificial ambient environments in greater detail.

3 Detection via Artificial Ambient Environments

As an alternative to the natural ambient environment, we proposed the generation of an Artificial Ambient Environment (AAE) using the peripherals of the transaction devices [12]. In this section, we discuss the basic principles of using AAEs as an anti-relay mechanism. Firstly, we present how infrared light can be used as an AAE actuator, before describing the use of sound as a proximity detection mechanism and as a communication medium for proximate devices. Finally, we suggest how other actuators may be used to provide proximity assurances.

Table 5. Estimated EER for machine learning algorithms, obtained by repeating 10-fold cross-validation 10 times.

Sensor	Random Forest	Naive Bayes	Decision Tree	Logistic Regression	Support Vector Machine
Accelerometer	**0.277** ± 0.052	0.474 ± 0.047	0.358 ± 0.059	0.483 ± 0.050	0.454 ± 0.126
Gyroscope	**0.179** ± 0.041	0.354 ± 0.059	0.228 ± 0.049	0.356 ± 0.055	0.288 ± 0.045
Magnetic field	**0.361** ± 0.055	0.400 ± 0.053	0.389 ± 0.063	0.421 ± 0.061	0.385 ± 0.053
Rotation vector	**0.285** ± 0.052	0.327 ± 0.055	0.317 ± 0.073	0.353 ± 0.050	0.325 ± 0.050
Gravity	0.499 ± 0.046	0.488 ± 0.043	0.494 ± 0.057	**0.484** ± 0.043	0.486 ± 0.156
Light	0.361 ± 0.059	0.369 ± 0.058	**0.293** ± 0.149	0.407 ± 0.054	0.351 ± 0.054
Linear acceleration	**0.307** ± 0.050	0.484 ± 0.048	0.392 ± 0.057	0.502 ± 0.049	0.397 ± 0.058

3.1 Artificial Ambient Environments

In order to increase the irreproducibility and uniqueness of an ambient environment, the transaction devices generate an artificial environment using their peripherals – measurable by a particular ambient sensor(s). The artificial environment should be based on randomly generated bits or sequences to act as a second (out-of-band) channel for assuring proximity between the transaction devices (see Fig. 5).

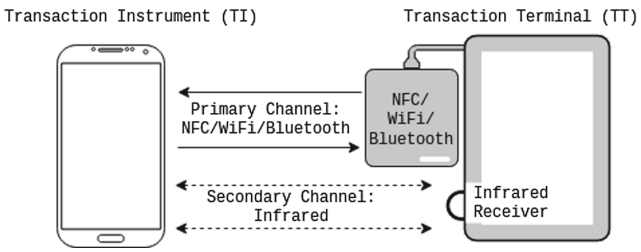

Fig. 5. High-level communication overview.

Upon initiation of a transaction, one (unidirectional) or both (bidirectional) device(s) are responsible for the generation and/or sensing of the AAE for some predefined time. After recording the sensor measurement, a comparison is performed with the captured data from both devices. The comparison may be performed by one of the communicating parties or by a trusted third party, as discussed in Sect. 2.1.

During the comparison, only the data that was captured while the artificial ambient channel was active should be considered; data captured outside this time-frame should be discarded. This way, an attacker cannot capture the generated sequence and then replay it at a remote location. Moreover, for an effective

AAE, the attacker should not be able to relay the data from the out-of-band channel in a way that the trusted comparison party cannot distinguish between a legitimate and an illegitimate transaction with a high degree of confidence.

To summarise, the basic principles of an AAE are:

1. The AAE generation should be based on random bits/sequences.
2. The AAE should provide sufficient evidence in order for two genuine devices to establish proximity assurance.
3. It should be hard for the attacker to accurately reproduce the AAE at a remote location.

The primary goal of the AAE is to protect against the off-the-shelf attacker. A resourceful attacker with access to state of the art equipment might be capable of effectively reproducing the same conditions at a remote location in a timely fashion. However, smartphones suffer from a plethora of security issues [26] and, in practice, a resourceful attacker is more likely to exploit these than invest in state of the art hardware to conduct a relay attack. On modern smartphones, widely-available peripherals that could potentially act as AAE actuators include: 1. the device's infrared emitter, 2. speaker, 3. flash light, 4. vibration, 5. display, 6. WiFi, 7. Bluetooth, 8. camera.

3.2 Infrared Light as an AAE Actuator

In [12], the use of infrared light as an AAE actuator was empirically evaluated. The AAE generation was based on 500 random bits, represented by 200 µs long pulses (1s) and pauses (0s) of the infrared emitter (therefore the total emission time was 100 ms). The bit sequence '1101110011', for instance, would be represented by the stream shown in Fig. 6.

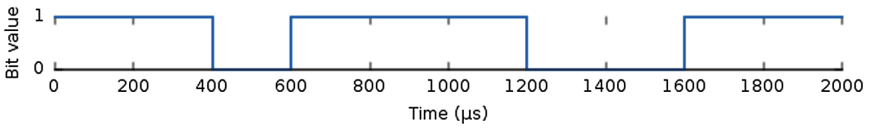

Fig. 6. Representation of the bit sequence '1101110011' in pulses-pauses.

The transaction instrument (TI) begins the infrared emission process when the transaction is initiated. The transaction terminal (TT) listens for infrared signals for 100 ms plus some acceptable offset window (4 ms), and rejects any signals received outside of this time-frame. Due to intrinsic hardware delays encountered during the experiments, TI was not able to immediately initiate the emission process, and some time x_i was required prior to the process to compensate for this. This time was quantifiable because the total emission time (100 ms) was known, as well as the total time required between the initiation and completion of the emission process. The bits accepted by device TT hence

depended upon time x_i. Any bits captured prior to ($x_i - 2$ ms) and after ($x_i + 100$ ms $+ 2$ ms) were rejected, where the 2 ms before and after comprises the acceptable offset. The offset is the maximum allowed deviation from the average time required by the transaction initiation. Figure 7 depicts the process on the two channels.

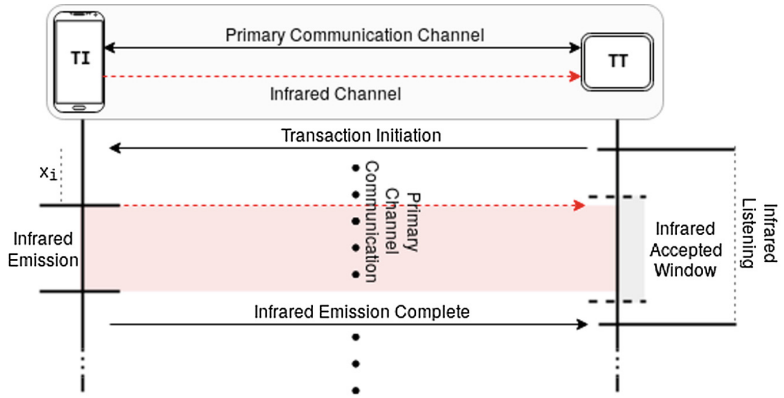

Fig. 7. Infrared as an AAE actuator.

The generated bits from TI and the captured bits from TT are compared for similarity after this process. As mentioned in Sect. 2.1, the comparison itself may take place on the terminal or by a trusted third party. Further investigations [13] showed that the process overhead is minimal, and the comparison can take place effectively during the transaction time (500 ms) by one of the devices.

The assumption of this technique is that an attacker, using off-the-shelf equipment, cannot effectively relay infrared data within that time-frame without being detected. A delay of more than 200 µs in relaying a single bit would introduce new bits into the sequence. Caching might increase the potential for an attacker to evade the proposed solution, since the risk of introducing extra bits is shifted from the bit level to the cache length level (when the relay is delayed by more than 200 µs). However, extensive caching (more than approximately 4 ms) would prolong the completion of the emission process beyond the acceptable time-limit of 100 ms, and so the delayed bits would not be considered. Hence, caching and subsequently relaying segments of the captured random sequence would have to be limited a maximum of a few milliseconds. Moreover, relay equipment such as fibre optics, where a cable connects the two relay devices, were not considered, as it can be easily detectable by the victim and/or terminal operator.

During a legitimate transaction, the similarity between the emitted and captured data was measured to be 98% or higher on approximately 98% of the performed transactions. This, therefore, was set as a baseline threshold. The

`dwdiff` tool[3] was used for the comparison of the two bit-streams. Six distinct relay test-beds were developed using off-the-shelf equipment, such as infrared extenders, Raspberry Pis, and mobile devices. None of the test-beds could effectively attack the proposed solution. The highest similarity rate after conducting a relay attack was achieved by an infrared extender, with 95% similarity across approximately 10% of all performed transactions. The reader is referred to [12] for a detailed discussion of the test-beds and the results.

3.3 Sound as an AAE Actuator

To use sound as an AAE actuator, one or both of the transaction devices generate and play a random sound through their speakers for some predefined time. In the event that only one device is playing the sound, the other should be recording. The captured sound-waves should then undergo a similarity comparison upon completing the recording. In case both devices are playing a randomly generated sound, they should also record simultaneously. The two captured sound-waves should again be compared against each other upon completion. A variation of this approach was investigated by Li et al. [22]. Even though the primary purpose was to restrict the communicating distance of two devices, they also demonstrate the effectiveness against relay attacks as part of their solution. In this work, the acoustic channel was used as the main channel for communicating messages between two devices. Signals were transmitted at a high frequency by both devices simultaneously (full-duplex communication). The full-duplex approach assisted towards relay attack prevention, as each device was capturing signals from both communicating devices, including itself. It could therefore estimate whether a message from the opposite party was received within some acceptable window by juxtaposing it with the message emanating from itself. Positive results were reported by the authors with a high success rate in silent environments; however, this degraded as the surrounding environment became noisier.

Karapanos et al. [21] explored the use of sound as a two-factor authentication method; however, attacks and potential solutions against this method have been demonstrated by Shrestha et al. [33].

Yi et al. [39] used the acoustic channel as a means of user authentication for unlocking mobile devices using a wearable device (smartwatch). The authors reported promising results with a lower bit error rate than conventional smartphone unlocking methods, e.g. PIN entry. While the main purpose of this work was not to counter relay attacks, they argue that such attacks would be expensive to carry out and, as the size of the relay equipment is large, it could easily be identified by a genuine user. They also claim that fingerprinting techniques can be used to uniquely identify the acoustic hardware to determine whether it originated from a genuine or relay device. Lastly, they mention that distance bounding protocols can be used, but a full investigation of this was considered out of scope.

[3] `dwdiff` tool: http://os.ghalkes.nl/dwdiff.html.

Based on the analysis in [31], while promising results have been demonstrated through the use of sound as a PRAD mechanism, it might not be applicable in EMV, transport ticketing, or other transactions with industry-imposed time restrictions of up to 500 ms. On average, due to latency related to initiating the recording process, recordings lasted for less than 280 ms within a 500 ms permitted time-frame. Similar latencies were not observed on most of the evaluated sensors; we concluded that sensor hardware may have a significant bearing on the effectiveness of a PRAD mechanism.

3.4 Other AAE Actuators

In this section, we discuss other potential AAE actuators; we focus on the candidate actuators listed in Sect. 3.1. In some cases, like in the case the Bluetooth or the WiFi, the underlying technology may not be flexible enough to be used as an AAE actuator without substantial modifications.

The display of a device could be used as a potential AAE actuator in combination with the camera of the communicating device. One device could display a randomly generated video feed to be recorded by the other device; the displayed and captured video feeds may then be compared for similarity. The advantage of this technique is that relaying video may incur a relatively large degree of latency. The main downsides are that: 1. the two devices ought to be held in the correct way to maximise success; and 2. the delay in capturing and playback of a stream could potentially negatively affect the results.

Similarly, the device's flashlight could be used by displaying a random pattern that is captured by the communicating party's camera or light sensor. Previous work has achieved reliable emission at speeds up to 500 bits per second (bps) using a LED flashlight and 15 bps using a Xenon flashlight [10,11,19], but this is 5 times slower than a typical infrared emitter, as per [12]. As such, an attacker would have a much larger relay window, which may hinder the effectiveness of this method. Additionally, the flashlight falls within the visible light spectrum, which may physically disturb nearby users.

One other potential option is vibration. While vibration has not been used to the best of our knowledge as a relay attack detection mechanism, it has been used to authenticate RFID tags [28] and to exchange secrets [29] with a high success rate. This evidence suggests that there is a potential in using vibration as an anti-relay mechanism. Here, one or both of the transaction devices generate a random vibration pattern, which is measured by both devices using a motion sensor, e.g. accelerometer. The captured data can be transmitted to a trusted third party upon the completion of the transaction for comparison.

4 Conclusion and Future Work

Proximity and Relay-Attack Detection (PRAD) is an important element for many contactless and wireless technologies. In this work, we illustrated that the viability of PRAD mechanisms can be largely dependent on the time constraints

mandated by industry requirements. Contactless payment transactions for example – whether smart card- or smartphone-based – must adhere to <3 cm for proximity and <500 ms for transaction duration, as stipulated by the EMV specifications. We evaluated the claim that natural ambient environments can provide a robust PRAD, as stated by some previous literature, under industry-specified time constraints. This was evaluated for both proximity detection (Sect. 2.4) and as a relay attack detection mechanism using a test-bed that reflected an actual attack (Sect. 2.5). We presented the results of a two-part evaluation using six similarity metrics used previously and several widely-used machine learning classifiers. In all cases, the results were far from what was claimed in past literature; our initial results indicate that natural ambient environments provide a poor PRAD for time-critical domains such as banking and transport. As such, we strongly recommend that any PRAD proposal should be evaluated based on the operating restrictions of the suggested deployment application.

This led to the development of artificially generated environments, which are random and unique to each transaction, for providing a more effective PRAD mechanism. To test this, we proposed a framework for deploying an artificial ambient environment (AAE) for PRAD. We developed a test-bed to evaluate the effectiveness of infrared in conjunction with six relay attack test-beds. In all cases, the genuine and relayed transactions were distinguishable for 97–98% of all transactions – far greater than the results using natural ambient sensing from our investigations. At present, we are expanding our interrogation of natural environment-based PRADs, using multiple sensors simultaneously with a range of sensor fusion techniques. Moreover, we are investigating the applicability of other smartphone sensors as an AAE-based PRAD mechanism. The first phase of this evaluation has been conducted using vibration as an AAE, which has yielded promising results.

Acknowledgement. Carlton Shepherd is supported by the EPSRC and the British government as part of the Centre for Doctoral Training in Cyber Security at Royal Holloway, University of London (EP/K035584/1). The authors would also like to thank anonymous reviewers for their valuable comments.

References

1. Transit and Contactless Open Payments: An Emerging Approach for Fare Collection. White paper, Smart Card Alliance Transportation Council, November 2011
2. How to Optimize the Consumer Contactless Experience? The Perfect Tap. Technical report, MasterCard (2014)
3. EMV Contactless Specifications for Payment Systems: Book D - EMV Contactless Communication Protocol Specification. Spec V2.6, EMVCo, LLC, March 2016
4. Transactions Acceptance Device Guide (TADG). Specification Version 3.1, VISA, November 2016
5. Boureanu, I., Mitrokotsa, A., Vaudenay, S.: Towards secure distance bounding. In: Moriai, S. (ed.) FSE 2013. LNCS, vol. 8424, pp. 55–67. Springer, Heidelberg (2014). doi:10.1007/978-3-662-43933-3_4

6. Coskun, V., Ozdenizci, B., Ok, K.: A survey on Near Field Communication (NFC) technology. Wireless Pers. Commun. **71**(3), 2259–2294 (2013). http://dx.doi.org/10.1007/s11277-012-0935-5
7. Cremers, C., Rasmussen, K., Schmidt, B., Capkun, S.: Distance hijacking attacks on distance bounding protocols. In: 2012 IEEE Symposium on Security and Privacy, pp. 113–127, May 2012
8. Francis, L., Hancke, G., Mayes, K., Markantonakis, K.: Practical NFC peer-to-peer relay attack using mobile phones. In: Ors Yalcin, S.B. (ed.) RFIDSec 2010. LNCS, vol. 6370, pp. 35–49. Springer, Heidelberg (2010). doi:10.1007/978-3-642-16822-2_4
9. Francis, L., Hancke, G.P., Mayes, K., Markantonakis, K.: Practical relay attack on contactless transactions by using NFC mobile phones. In: IACR Cryptology Archive 2011, p. 618 (2011)
10. Galal, M.M., Fayed, H.A., Aziz, A.A.E., Aly, M.H.: Smartphones for payments and withdrawals utilizing embedded LED flashlight for high speed data transmission. In: 2013 Fifth International Conference on Computational Intelligence, Communication Systems and Networks, pp. 63–66, June 2013
11. Galal, M.M., Aziz, A.A.A.E., Fayed, H.A., Aly, M.H.: Smartphone payment via flashlight: utilizing the built-in flashlight of smartphones as replacement for magnetic cards. Optik - Int. J. Light Electron Optics **127**(5), 2453–2460 (2016)
12. Gurulian, I., Akram, R.N., Markantonakis, K., Mayes, K.: Preventing relay attacks in mobile transactions using infrared light. In: Proceedings of the Symposium on Applied Computing, SAC 2017, pp. 1724–1731. ACM, New York (2017)
13. Gurulian, I., Markantonakis, K., Akram, R.N., Mayes, K.: Artificial ambient environments for proximity critical applications. In: 2017 12th International Conference on Availability, Reliability and Security, ARES 2017. ACM, New York (2017)
14. Gurulian, I., Shepherd, C., Frank, E., Markantonakis, K., Akram, R., Mayes, K.: On the effectiveness of ambient sensing for NFC-based proximity detection by applying relay attack data. In: The 16th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2017. IEEE, August 2017
15. Haken, G., Markantonakis, K., Gurulian, I., Shepherd, C., Akram, R.N.: Evaluation of Apple iDevice sensors as a potential relay attack countermeasure for Apple Pay. In: Proceedings of the 3rd ACM Workshop on Cyber-Physical System Security, CPSS 2017, pp. 21–32. ACM, New York (2017)
16. Halevi, T., Ma, D., Saxena, N., Xiang, T.: Secure proximity detection for NFC devices based on ambient sensor data. In: Foresti, S., Yung, M., Martinelli, F. (eds.) ESORICS 2012. LNCS, vol. 7459, pp. 379–396. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33167-1_22
17. Hancke, G.P., Kuhn, M.G.: Attacks on time-of-flight distance bounding channels. In: Proceedings of the First ACM Conference on Wireless Network Security, WiSec 2008, pp. 194–202. ACM, New York (2008). http://doi.acm.org/10.1145/1352533.1352566
18. Hancke, G., Mayes, K., Markantonakis, K.: Confidence in smart token proximity: relay attacks revisited. Comput. Secur. **28**(7), 615–627 (2009). http://www.sciencedirect.com/science/article/pii/S0167404809000595
19. Hesselmann, T., Henze, N., Boll, S.: FlashLight: optical communication between mobile phones and interactive tabletops. In: ACM International Conference on Interactive Tabletops and Surfaces, ITS 2010, pp. 135–138. ACM, New York (2010), http://doi.acm.org/10.1145/1936652.1936679

20. Jin, R., Shi, L., Zeng, K., Pande, A., Mohapatra, P.: MagPairing: pairing smartphones in close proximity using magnetometers. IEEE Trans. Inf. Forensics Secur. **11**(6), 1306–1320 (2016)
21. Karapanos, N., Marforio, C., Soriente, C., Capkun, S.: Sound-Proof: usable two-factor authentication based on ambient sound. In: 24th USENIX Security Symposium. USENIX Association, Washington, D.C., August 2015
22. Li, L., Xue, G., Zhao, X.: The power of whispering: near field assertions via acoustic communications. In: Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security, ASIA CCS 2015, pp. 627–632. ACM, New York (2015). http://doi.acm.org/10.1145/2714576.2714586
23. Ma, D., Saxena, N., Xiang, T., Zhu, Y.: Location-aware and safer cards: enhancing RFID security and privacy via location sensing. IEEE TDSC **10**(2), 57–69 (2013)
24. Maltoni, D., Maio, D., Jain, A., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer Science & Business Media, London (2009). doi:10.1007/978-1-84882-254-2
25. Mehrnezhad, M., Hao, F., Shahandashti, S.F.: Tap-Tap and Pay (TTP): preventing man-in-the-middle attacks in NFC payment using mobile sensors. In: 2nd International Conference on Research in Security Standardisation, October 2014
26. Polla, M.L., Martinelli, F., Sgandurra, D.: A survey on security for mobile devices. IEEE Commun. Surv. Tutorials **15**(1), 446–471 (2013)
27. Rasmussen, K.B., Capkun, S.: Realization of RF distance bounding. In: USENIX Security Symposium, pp. 389–402 (2010)
28. Saxena, N., Uddin, M.B., Voris, J., Asokan, N.: Vibrate-to-unlock: mobile phone assisted user authentication to multiple personal RFID tags. In: 2011 IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 181–188, March 2011
29. Shen, Z., Zheng, X., Xie, H.: Near field service initiation via vibration channel. In: 2016 12th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN), pp. 450–453, December 2016
30. Shepherd, C., Akram, R.N., Markantonakis, K.: Towards trusted execution of multi-modal continuous authentication schemes. In: Proceedings of the 32nd Symposium on Applied Computing, pp. 1444–1451. ACM (2017)
31. Shepherd, C., Gurulian, I., Frank, E., Markantonakis, K., Akram, R., Mayes, K., Panaousis, E.: The applicability of ambient sensors as proximity evidence for NFC transactions. In: Mobile Security Technologies, IEEE Security and Privacy Workshops, MoST 2017. IEEE, May 2017
32. Shrestha, B., Saxena, N., Truong, H.T.T., Asokan, N.: Drone to the rescue: relay-resilient authentication using ambient multi-sensing. In: Christin, N., Safavi-Naini, R. (eds.) FC 2014. LNCS, vol. 8437, pp. 349–364. Springer, Heidelberg (2014). doi:10.1007/978-3-662-45472-5_23
33. Shrestha, B., Shirvanian, M., Shrestha, P., Saxena, N.: The sounds of the phones: dangers of zero-effort second factor login based on ambient audio. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS 2016 pp. 908–919. ACM, New York (2016)
34. Truong, H.T.T., Gao, X., Shrestha, B., Saxena, N., Asokan, N., Nurmi, P.: Comparing and fusing different sensor modalities for relay attack resistance in zero-interaction authentication. In: 2014 IEEE International Conference on Pervasive Computing and Communications, pp. 163–171. IEEE (2014)
35. Umar, A., Mayes, K., Markantonakis, K.: Performance variation in host-based card emulation compared to a hardware security element. In: 2015 First Conference on Mobile and Secure Services, pp. 1–6. IEEE (2015)

36. Urien, P., Piramuthu, S.: Elliptic curve-based RFID/NFC authentication with temperature sensor input for relay attacks. Decision Support Syst. **59**, 28–36 (2014)
37. Varshavsky, A., Scannell, A., LaMarca, A., de Lara, E.: Amigo: proximity-based authentication of mobile devices. In: Krumm, J., Abowd, G.D., Seneviratne, A., Strang, T. (eds.) UbiComp 2007. LNCS, vol. 4717, pp. 253–270. Springer, Heidelberg (2007). doi:10.1007/978-3-540-74853-3_15
38. Verdult, R., Kooman, F.: Practical attacks on NFC enabled cell phones. In: 2011 3rd International Workshop on Near Field Communication (NFC), pp. 77–82, February 2011
39. Yi, S., Qin, Z., Carter, N., Li, Q.: WearLock: unlocking your phone via acoustics using smartwatch. In: 2017 IEEE 37th IEEE International Conference on Distributed Computing Systems, ICDCS 2017 (2017)