Research Article

Attack Potential Evaluation in Desktop and Smartphone Fingerprint Sensors: Can They Be Attacked by Anyone?

Ines Goicoechea-Telleria , **Raul Sanchez-Reillo,**
Judith Liu-Jimenez, and Ramon Blanco-Gonzalo

The University Group for Identification Technologies, Carlos III University of Madrid, 28911 Leganes, Spain

Correspondence should be addressed to Ines Goicoechea-Telleria; igoicoec@ing.uc3m.es

Received 30 November 2017; Revised 22 January 2018; Accepted 1 February 2018; Published 16 April 2018

Academic Editor: Ding Wang

Copyright © 2018 Ines Goicoechea-Telleria et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The use of biometrics keeps growing. Every day, we use biometric recognition to unlock our phones or to have access to places such as the gym or the office, so we rely on the security manufacturers offer when protecting our privileges and private life. It is well known that it is possible to hack into a fingerprint sensor using fake fingers made of Play-Doh and other easy-to-obtain materials but to what extent? Is this true for all users or only for specialists with a deep knowledge on biometrics? Are smartphone fingerprint sensors as reliable as desktop sensors? To answer these questions, we performed 3 separate evaluations. First, we evaluated 4 desktop fingerprint sensors of different technologies by attacking them with 7 different fake finger materials. All of them were successfully attacked by an experienced attacker. Secondly, we carried out a similar test on 5 smartphones with embedded sensors using the most successful materials, which also hacked the 5 sensors. Lastly, we gathered 15 simulated attackers with no background in biometrics to create fake fingers of several materials, and they had one week to attack the fingerprint sensors of the same 5 smartphones, with the starting point of a short video with the techniques to create them. All 5 smartphones were successfully attacked by an inexperienced attacker. This paper will provide the results achieved, as well as an analysis on the attack potential of every case. All results are given following the metrics of the standard ISO/IEC 30107-3.

1. Introduction

Biometric recognition has become a commonplace technology in our everyday lives. We use it to unlock our phones, to get in the gym or to enter the office due to its comfort of use [1] and freeing us from remembering passwords [2, 3]. Nevertheless, when sensitive data (personal pictures, documents) or privileges (gym membership, food coupons) are at stake, we need to make sure that we know how secure these systems are.

There have been many tests of many fingerprint sensors' ability to detect attacks in the last decades [4–12]. For that end, different materials were used to create fake fingers and use them on sensors, to check if an attacker would be able to bypass the security. When the first smartphones with embedded fingerprint sensors were released, fake fingers were created to try to attack the system and succeeded [13]. It was rapidly spread on the media that biometric recognition was not secure and that people should not trust it. Soon after, it

was proven that an attacker can steal a person's fingerprint by taking a picture of it from a distance [14].

The simulated attackers that performed these tests were, to the best of our knowledge, researchers or proficient in biometrics. Moreover, the released videos only show one attempt, which showed to be successful. However, how many attempts did they need until it worked? How long had they been working on biometrics? What was their expertise? These questions were left unanswered.

To answer them, it is necessary to follow a common standard to have comparable security evaluations and to give a complete understanding of how the systems behave against attacks. This is what was achieved with this work. To fulfill the need of comparable security evaluations, there are several tools like Common Criteria [15] and its evaluation methodology, CEM (Common Methodology for Information Technology Security Evaluation) [16]. These are focused on Information Technology security in general and need some adaptation for the case of biometrics, and more particularly

for the attacks at the sensor capture level, the so-called presentation attacks. Hence, a new ISO standard was created, ISO/IEC 30107 [17], to address this need for Presentation Attack Detection (PAD) evaluation standardization. Also, some works have been done on methodologies and best practices to evaluate security [18, 19] and to evaluate the performance of sensors embedded in smartphones [20]. A methodology unifying both was proposed in [21]. This work gathers 3 studies following thoroughly the said standards, in order to give a complete answer to the attack resistance ability of these systems.

First, an experiment was made with 4 desktop fingerprint sensors: 1 thermal, 2 capacitive, and 1 optical. In total, 4,672 attacks were attempted using 7 different artefact species (fake finger material) [22]. Both cooperative and noncooperative tests were made. Cooperative attacks require the collaboration of the capture subject to get their biometric characteristic on a mold, while, for noncooperative attacks, the attacker gets the fingerprint from a surface without the collaboration of the capture subject.

For the second study, the 3 most interesting artefact species (according to ease of production, success on attacks, or level of resemblance to real finger) from the first study were chosen to attempt attacks on 5 smartphones that have embedded fingerprint sensors, summing a total of 2,669 attacks [21]. This experiment was performed by the same evaluator of the first experiment, thus having gathered knowledge on how to perform attacks. All the artefacts were created in a collaborative manner in this case.

Finally, for the third study, we gave 15 simulated attackers with no background in biometrics one week to attack one smartphone's fingerprint sensor (5 smartphones in total, same ones than in the second study). Each had to, at least, use 3 bona fide capture subjects (but they could use more for extra credit) and use each material at least 120 times on the smartphone sensor, making a total of 5,841 attempts. As more than one week would be needed to create noncooperative fake fingers, the study was focused only on cooperative attacks.

With these 3 studies following the same standard (ISO/IEC 30107) and methodology, it is possible to make a comparison of the attack potential for each case and get results and thus answer if attacking fingerprint sensors is a matter of expertise and how many attempts are needed to successfully attack a sensor, in average.

This paper is divided into 6 sections. Section 1 includes an introduction and Section 2 gives an overview on related work. Section 3 is focused on the methodology carried out during the evaluations and Section 4 analyzes Common Criteria's attack potential for each case. Results are reported and discussed in Section 5 and Section 6 describes the conclusions that were obtained during this work, as well as future work.

2. Related Work

There can be many vulnerable points in a biometric system: at presentation level, identity claim, data transfer, quality and feature extraction, decision thresholds, and so on [23]; and this also applies to mobile devices. For instance, some

vulnerabilities were found creating a malicious application that steals the temporary fingerprint image by accessing its memory space or extracting a stored template from the nonvolatile memory and recreating the feature points of the fingerprint [24]. In [25, 26], privacy issues are addressed for biometric user authentication, and they propose some countermeasures for a properly designed secure and privacy-preserving system.

In addition, several security analyses have been made using altered fingerprints [27, 28] and one was performed specifically on mobile devices [29]. Nevertheless, it differs slightly from the topic of this paper, as it is focused only on alterations to fingerprint sample images, feeding them directly to the system, and not on creating artefacts and molds.

Many studies and evaluations have been performed regarding presentation attacks on fingerprint sensors. Already in 1990, several sensors started being tested using artefacts, and the system failed to reject them even from the first attempt [4, p. 15]. On 2000, an evaluation was performed on [10, p. 9] by calculating the acceptance rate of 1 user's finger made with gelatin on 11 sensors, where the artefacts were accepted by the systems in a very high percentage (the lowest being 67% fake finger acceptance rate). On 2002, several more attacks were demonstrated by using latent fingerprint reactivation on 6 capacitive, 2 optical, and 1 thermal scanners [30]. For the case of [7], 10 subjects were used to create gelatin fingers and use them on 3 sensors, getting success rates from 44.6% to 76% success rates. In all experiments, only index fingers were used. In general, nevertheless, these studies do not follow a thorough evaluation procedure nor standard and merely prove when a certain material or technique is effective on specific sensors at least once.

In 2009 [31], Liveness Detection (*LivDet*) competitions started and continued in 2011 [32], 2013 [33], and 2015 [34]. Their goal was to compare different liveness detection (Presentation Attack Detection) mechanisms by using them on a very large database of fake fingers (made of gelatin, latex, ecoflex, Play-Doh, silicone, wood glue, and modasil). Different academic institutions or industries could try their algorithms on the database. Four different sensors were used to acquire the images and the evaluations were done using a common testing protocol.

To the authors' knowledge, there have not been evaluations specifically focused on attacking mobile devices with fake fingers, but there have been reports on found vulnerabilities. In 2013, when the first iPhone with an integrated fingerprint sensor came out, the Chaos Computer Club [13] proved that it was possible to break into the sensor using a white glue fake finger covered with graphite, and the fingerprint could be stolen from the phone screen using a scanner and doing some image processing. Nonetheless, this was only reported once in a video, not in a complete evaluation. In 2016, fake fingers were printed using conductive ink (having a sample of the fingerprint image beforehand), so they could be used directly on the mobile phone sensor without having to create molds previously [35]. This was a technical report to inform about the vulnerability and not an evaluation.

There are several ways to overcome these attacks, divided into two main groups: software and hardware PAD mechanisms [5]. Software PAD mechanisms take the sample captured by the sensor and apply image processing and classification techniques to tell whether the finger is real or not. On the other hand, hardware PAD adds additional sensors (temperature sensor, multispectral cameras, etc. [36]) to make this distinction. Hardware solutions have lower error rates than the software ones [37] but are usually more expensive or bulky due to the additional equipment needed [36, p. 10]. Thus, hardware solutions are usually not included in mobile devices that should be cheaper and smaller.

3. Materials and Methods

The methodology was as homogeneous as possible across all studies, following evaluation methods from Common Criteria's CEM [16] and ISO/IEC 30107 [17], which were unified as a methodology on [21]. As for every biometrics evaluation, three steps are needed: planning, execution, and results reporting. In this case, the results will be given later in Section 5.

3.1. Planning the Evaluation. In the case of the desktop fingerprint sensors, an expert performed an evaluation on 1 thermal, 2 capacitive, and 1 optical sensors. The process used to create the artefacts was the usual one seen in many evaluations and research on PAD [10, pp. 5–8], both cooperative (capture subject cooperates in the creation of the mold) and noncooperative (attacker steals biometric characteristic with no help from capture subject). For this study, it was important to try many artefact species (i.e., Play-Doh, gelatin, latex, silicone, white glue, latex with graphite, and silicone with graphite) to check which ones were more threatening to the systems.

For the second study, an expert performed an evaluation on 5 mobile devices with an embedded fingerprint sensor. The artefact species (materials) that turned out to be most threatening for the previous evaluation were used for this one: Play-Doh, gelatin, and latex with graphite.

Lastly, for the case of the third study, we gathered 15 simulated attackers with no background in attacking biometric systems. We prepared a process or “recipe” they had to follow to create the artefacts and we gave it to them on writing and video. They had to create artefacts with Play-Doh and gelatin but could get extra points for using more artefact species. Again, there were 5 smartphones available for the evaluation (same ones as in study 1) and each simulated attacker was given one at random. Also, those who owned one of these cell phones could use them for the attack. Thus, some smartphones were used more than others. They had one week to perform at least 120 attack attempts per artefact species. To make sure that they were performing the evaluation correctly, attackers had to record themselves on video using the fake fingers on the sensors. Moreover, they had to take pictures of all molds and artefacts and finally had to hand over a box with all of them.

TABLE 1: Characteristics of desktop fingerprint sensors according to sensor interaction type and sensing technology.

Fingerprint sensor	Sensor type	Technology
THER	Touch	Thermal
CAP1	Touch	Capacitive
CAP2	Touch	Capacitive
OPT	Touch	Optical

TABLE 2: Characteristics of the mobile device sensors used for the evaluation, according to sensor type, shape, and location on the device.

Mobile device (MD)	Sensor type	Sensor shape	Sensor location
MD1	Swipe	Rectangular	Front
MD2	Touch	Circular	Front
MD3	Touch	Rectangular	Front
MD4	Touch	Rectangular	Side
MD5	Touch	Circular	Back

FIGURE 1: Fingerprint sensor placement on smartphones.

On the following subsections, the TOE (Target of Evaluation) and target application will be described for each study and, with these in mind, the penetration test will be specified.

3.1.1. Description of the TOE (Target of Evaluation). Four fingerprint sensors with different sensing technologies were used for the first study. All of them were a gray box, because the only intermediate result that could be obtained was the quality score, measured by NIST's NFIQ quality algorithm [38]. The brands cannot be disclosed, but their main characteristics can be seen in Table 1.

Both PAD evaluations on mobile devices were performed on 5 smartphones with different embedded fingerprint sensors. Their main characteristics (sensor type, shape, and location) can be seen in Table 2 and Figure 1.

As the biometric systems evaluated on these 2 studies are full systems embedded in mobile devices, they are a black box that only reveals whether the verification with the artefact passed or failed.

3.1.2. Description of Target Application. A PAD evaluation has no meaning unless the target evaluation and the conditions

under which it is performed are specified. These differ in the case of desktop sensors and smartphones.

(a) *Desktop Fingerprint Sensors*. In the case of desktop fingerprint sensors, there are many possible applications, for example, entering an office, spending food coupons at work, or entering critical infrastructures (factories and nuclear plants). Thus, the consequences of an attacker hacking those sensors would be having an unauthorized person enter a certain building, which could be more or less critical, or having access to privileges that do not belong to the person.

The implemented functions for the systems are enrolment and verification, being their policies in detail:

- (i) Enrolment: all images of the artefacts were captured in an acquisition process. There were 2 transactions to obtain images with an NFIQ (NIST Fingerprint Image Quality) value equal or lower than 3 (good quality), and 3 attempts were made for each transaction until both images were successfully compared. If after all the attempts, no successful comparisons could be done between the two samples, the enrolment for this particular finger could not be completed.
- (ii) Verification: the verification was done offline. The artefacts were captured in the same enrolment process as the real fingers, so they could be discarded by NFIQ higher than 3 (low quality). Then, in an offline process, the artefact samples were compared with the NBIS algorithm to the real fingers.

(b) *Mobile Phones*. The target application in the case of smartphones is to unlock the user's mobile phone. From there, all apps can be accessed (as only a few require additional security, like a PIN), including bank accounts, password managers, or personal pictures. Moreover, some apps have implemented fingerprint logging-in by using the system's fingerprint manager. For the case of smartphones, there is never any surveillance by a guard (as it is used solely by the owner). The implemented functions in this case are also enrolment and verification, with the particularity that the evaluator cannot decide on the policies.

- (i) Enrolment: it is crucial for the evaluation, as it can influence the performance, and each mobile device has a different enrolment policy. As it is a black box, the evaluator cannot decide on the enrolment and verification policies, as can be seen in Table 3.
- (ii) MD5's policy must be noted, as it only needs 6 samples to do the enrolment and that could influence the final performance of the fingerprint sensor.
- (iii) Verification: the artefacts are used in a verification process, that is, the artefact attempts to be verified as the real finger that has been previously enrolled. The number of allowed attempts for each smartphone is detailed on Table 4.

As it can be seen, devices MD1, MD3, MD4, and MD5 accept an unlimited number of attempts to attack them, as when 5 attempts have failed, they just wait for 30 seconds and

TABLE 3: Enrolment policy for each mobile device. Policies are given by the manufacturer and cannot be changed by the evaluator.

Mobile device	Attempt policy for enrolment
MD1	Around 10 attempts needed (can be extended 10 more times if the user wants). Asks the user to move finger to get different samples and <i>checks it</i> .
MD2	Around 15 attempts needed (10 for center part of the finger and 5 for corners). Asks the user to move finger to get different samples and <i>checks it</i> .
MD3	Around 20 attempts needed. Asks the user to move finger to get different samples and <i>checks it</i> .
MD4	Around 20 attempts needed. Asks the user to move finger to get different samples and <i>checks it</i> .
MD5	Around 6 attempts needed. It just asks the user to move the finger to get different samples, but <i>does not check it</i> .

TABLE 4: Transaction policies for verification for each mobile device. Policies are given by the manufacturer and cannot be changed by the evaluator.

Mobile device	Allowed failed attempts	Policy
MD1	5	Waits for 30 seconds, can do attack again
MD2	3	Asks for PIN
MD3	5	Waits for 30 seconds, can do attack again
MD4	5	Waits for 30 seconds, can do attack again
MD5	5	Waits for 30 seconds, can do attack again

the attacker can try again, with no more restrictions (as far as the author's knowledge). MD2 asks for a PIN after 3 attempts.

Also, for all cases, if the phone is turned off and the attacker wants to turn it on and access its data, he or she will need additional information apart from the bona fide user's fingerprint, like a PIN or a password. So, if the phone is found turned off or without power and the attacker does not know the user's additional information, he or she will not be able to gain access.

3.1.3. Specification of Penetration Test. Once the systems under test have been analyzed and described, it is necessary to specify the penetration test and how it will be performed. In Table 5, the characteristics of the test are shown, as well as the final amount of attempts for clarity purposes.

For the experiment of the desktop sensors, an app was developed for the capture process. As it can be seen in Figure 2, the program showed which finger to capture, the amount of attempts, and samples left and an image of the captured fingerprint. It allowed an enrolment and verification process.

For the experiment of the smartphones, a mobile app was made for iOS and Android (Figure 3). For acquiring the data, the visit screen is filled in by the evaluator (genuine user's ID, attacker's ID, device, finger ID, type of attack, and artefact species) and the app logs whether the attack succeeded or not. The enrolment was performed using the phone's native settings procedure.

TABLE 5: PAD evaluation characteristics for each study: desktop sensors, mobile devices (1 expert), and mobile devices (15 laymen). Details are given according to ISO/IEC 30107-3 requirements.

	Desktop sensors	Mobile devices (1 expert)	Mobile devices (15 laymen)	
Capture subjects	6	7 (3 female + 4 male)	3 (more for extra points)	53
Sources for artefacts	36 fingers, 6 capture subjects (index, middle and thumb, both hands)	42 fingers, 6 capture subjects (index, middle and thumb, both hands)	4 per capture subject (both thumbs, indexes)	212
Number of attempts	3 per transaction, 2 transactions per finger, 4 sensors (4672 in total)	10 per finger, 5 mobile devices (2669 in total)	10 per finger	5,481
Evaluator's access	Gray box (pass/fail result + NFIQ quality).	Black box (only pass/fail result). Very slight quality feedback on some devices (MD1 and MD3) ("finger too wet").		
Artefact species	7 (Play-Doh, gelatin, latex, silicone, white glue, latex with graphite, silicone with graphite)	3 (Play-Doh, gelatin, latex with graphite)	2 (Play-Doh and gelatin), more for extra credit	10 (Play-Doh, gelatin, gelatin with glycerin, silicone, TAC adhesive, wax, clay, wood glue, conductive ink, latex)
Attack type	Cooperative and non-cooperative	Cooperative

FIGURE 2: Example display of the desktop capture program.

3.2. Executing the Evaluation. After the careful planning of the evaluation, it was finally executed with the specified penetration test. The 3 steps needed for the execution according to CEM [16] are detection, capture, and processing.

3.2.1. Detection. Before the actual execution, the different artefact species were put to the test. If the sensor could detect the fingerprint, then it could be selected for the evaluation. Moreover, for the third study, simulated attackers tried more materials apart from Play-Doh and gelatin but found that not all of them were suitable.

3.2.2. Capture. As checking the quality is not always possible for the evaluator, some artefacts with different qualities (examined by the evaluator) can be used to check which ones are obtained successfully by the sensor and continue with that technique [39]. For desktop sensors, the quality of the sample could be measured with NFIQ and an image of the sample could be seen at the moment of capture. Thus, it was

FIGURE 3: Smartphone app (Android and iOS) for logging the PAD evaluation. A pass/fail result is logged for each attempt.

possible for the evaluator to improve the attack during the evaluation. In the case of mobile devices, some of them give a slight feedback on quality by telling the user that, for example, the finger is too wet. If an attacker is using a gelatin artefact and the smartphone prompts "finger too wet," the attacker will make another artefact with less proportion of water and try again.

Lastly, it must be noted that each mobile device will have different algorithms, sensor technologies, and quality and decision thresholds.

3.2.3. Processing. In the case of desktop fingerprint sensors, all images were stored for an offline verification with the database of subjects' real fingers by using the NBIS algorithm by NIST [40]. The most significant values from the analysis

were the proportion of times that an artefact was verified as a normal presentation and the proportion of images that were rejected by the system due to low quality ($\text{NFIQ} > 4$).

In the case of smartphones, no quality or similarity scores can be obtained, just pass/fail results. Thus, the data used for the posterior analysis was based solely on this.

4. Analysis of Attack Potential

The attack potential is a standardized measure given by Common Criteria. According to Common Criteria methodology [16], the attack potential is a measure of the *effort to be expended in attacking a TOE (Target of Evaluation) with a PAI (Presentation Attack Instrument), expressed in terms of an attacker's expertise, resources, and motivation*, which can be divided into more specific parameters. Thus, TOEs are given a rating to assess their resistance to specific attacks.

4.1. Threats and Attacks. Every threat has a corresponding possible attack [18] and they are analyzed before calculating the attack potential.

4.1.1. Desktop Fingerprint Sensors

(i) *Possible Threats.* Although fingerprint sensor systems can have vulnerabilities at many points, this study only focuses on the presentation attack side, that is, using an artefact generated from a user's real finger on the sensor. The intended operation of the system depends on the target application, for example, opening a door to an office or to a gym or getting privileges like food coupons.

(ii) *Possible Attacks.* The attack that can exploit the threat explained in the previous point is the presentation attack. The biometric characteristic can be obtained in two ways: with or without cooperation from the capture subject. In this case, both cooperative and noncooperative attacks were done. The level of expertise of the evaluator is proficient, although the materials needed for the evaluation can be found at any supermarket.

4.1.2. Mobile Phones

(i) *Possible Threats.* Fingerprint sensors embedded in mobile devices can have vulnerabilities at many points, too. The intended operation of the system is to unlock a smartphone, thus accessing private data. As it was said above, the only vulnerable point used for the scope of this paper is the capture process.

(ii) *Possible Attacks.* The attack that can exploit the threat from the first point is the presentation attack. In this case, only cooperative attacks were done. The level of expertise of the simulated attackers is low: they had no prior knowledge on how to attack fingerprint sensors.

4.2. Attack Potential Calculation. This calculation is used by the evaluator to determine whether or not the TOE is resistant to attacks assuming a specific attack potential of an attacker

[16, pp. 422–432]. If the evaluator determines that a potential vulnerability is exploitable in the fingerprint sensor, they must confirm that it is exploitable by doing penetration tests (as specified on Section 3.1.3).

With this in mind, the evaluator determines the minimum attack potential required by an attacker to successfully carry on an attack and arrives at some conclusion about the TOE's resistance to attacks. This attack potential is confirmed on the penetration tests performed in this work on Section 5.2.

A score can be assigned to each of the attack potential parameters following Common Criteria's CEM (Common Evaluation Methodology) [16, p. 429]. By adding all the values from the different parameters, the attack potential of an artefact species is rated as basic, enhanced-basic, moderate, high, or beyond high.

4.2.1. Desktop Fingerprint Sensors. Following CEM specifications, a score was given to every parameter of the attack potential to calculate its total rating (each score is given according to the table on [16, p. 429]). The attack potential will be different for cooperative and noncooperative attacks, being the expertise and the elapsed time the most differentiating factors.

(a) *Cooperative Attacks.* See Table 6.

(b) *Noncooperative Attacks.* See Table 7.

4.2.2. Mobile Phones. In this case, only cooperative attacks were made, and the attack potential will be the same for both mobile device studies.

As it was calculated in Tables 6, 7, and 8, the rating for cooperative attacks on desktop fingerprint sensors is 6.5 (basic) and 10 (enhanced-basic) for the noncooperative. For the case of mobile phones, the rating is 4.5 (basic). These scores are specified on CEM. Thus, the attacks would have to be considered in penetration testing for all evaluations assuming, respectively, minimum, basic, and minimum attack potentials (or higher). If penetration tests show that the attack is successful, the TOE would fail to resist against that attack potential. Further details on how to calculate the attack potential are given on Common Criteria's CEM [16, pp. 422–432].

5. Results and Discussion

Lastly, after executing the 3 separate evaluations, the obtained results are analyzed and compared in this section.

5.1. Metrics. The standard ISO/IEC JTC1 30107-3 requires specific metrics for PAD evaluation reporting. As the access to the system differs in desktop sensors and smartphones, Figure 4 shows which metrics are suitable for each case. The possible metrics are as follows:

- (i) *APNRR (attack presentation nonresponse rate):* proportion of attack presentations using the same PAI species that cause no response at the PAD subsystem or data capture subsystem

TABLE 6: Attack potential calculation for cooperative attacks on desktop fingerprint sensors. Scores assigned according to the classification from Common Criteria [16, p. 429].

	Preparation phase	PAI construction + exercising phase	Attack execution phase	Total factor rating	Score
Elapsed time	<1 day (capture subject is cooperative)	<1 day or <1 week (different material difficulty)	Few seconds (perform attack)	<1 week or <2 weeks	1.5
Expertise	Layman (materials can be obtained at normal stores)	Layman (easy to create)	Layman (not much expertise needed)	Layman	0
Knowledge of TOE	Public (well known on the internet that it works)	Public (manuals can be found on the internet)	Public (no knowledge needed)	Public	0
Window of opportunity	Unnecessary (no access to TOE needed)	Easy (access to TOE for practicing)	Easy (high chance the PAI will work)	Easy	1
Equipment	Standard (no equipment needed)	Specialized (some sensors might be hard to obtain)	Standard (no equipment needed)	Specialized	4
Overall attack rating	6.5 (Basic)				
Attack resistance	Minimum				

TABLE 7: Attack potential calculation for noncooperative attacks on desktop fingerprint sensors. Scores assigned according to the classification from Common Criteria [16, p. 429].

	Preparation phase	PAI construction + exercising phase	Attack execution phase	Total factor rating	Score
Elapsed time	<1 week (capture subject is noncooperative)	1 week (creating PAIs)	Few seconds (perform attack)	<2 weeks	2
Expertise	Layman (materials can be obtained at normal stores)	Proficient (process needs many steps)	Layman (not much expertise needed)	Proficient	3
Knowledge of TOE	Public (well known on the internet that it works)	Public (manuals can be found on the internet)	Public (no knowledge needed)	Public	0
Window of opportunity	Unnecessary (no access to TOE needed)	Easy (access to TOE for practicing)	Easy (high chance the PAI will work)	Easy	1
Equipment	Standard (no equipment needed)	Specialized (some sensors might be hard to obtain)	Standard (no equipment needed)	Specialized	4
Overall attack rating	10 (Enhanced-basic)				
Attack resistance	Basic				

- (ii) *APAR (attack presentation acquisition rate)*: proportion of attack presentations using the same PAI species from which the data capture subsystem acquires a biometric sample of sufficient quality
- (iii) *IAPMR (impostor attack presentation match rate)*: proportion of impostor attack presentations using the same PAI species in which the target reference is

matched. When it is not matched, IAPNMR is used (*Impostor Attack Presentation Nonmatch Rate*).

5.2. Penetration Test Results. With the metrics described on the last subsection and the schema from Figure 4, we built the graphs for the error rates of desktop and smartphone fingerprint sensors.

TABLE 8: Attack potential calculation for cooperative attacks on smartphone fingerprint sensors. Scores assigned according to the classification from Common Criteria [16, p. 429].

	Preparation phase	PAI construction + exercising phase	Attack execution phase	Total factor rating	Score
Elapsed time	<1 day (capture subject is cooperative)	<1 day or <1 week (different material difficulty)	Few seconds (perform attack)	<1 week or <2 weeks	1.5
Expertise	Layman (materials can be obtained at normal stores)	Layman (easy to create)	Layman (not much expertise needed)	Layman	0
Knowledge of TOE	Public (well known on the internet that it works)	Public (manuals can be found on the internet)	Public (no knowledge needed)	Public	0
Window of opportunity	Unnecessary (no access to TOE needed)	Easy (access to TOE for practicing)	Easy (high chance the PAI will work)	Easy	1
Equipment	Standard (no equipment needed)	Standard (but it is necessary to buy the TOE, which can be expensive)	Standard (no equipment needed)	Standard	2
Overall attack rating	4.5 (Basic)				
Attack resistance	Minimum

FIGURE 4: Metrics used for each case. Desktop fingerprint sensors give us more intermediate decisions than mobile devices.

5.2.1. Desktop Fingerprint Sensors. As these systems gave us feedback on IAPMR, APAR, and APNRR, they are represented in Figures 5 and 6.

The most meaningful value is IAPMR, as it shows the proportion of presentation attacks that defeat a comparison system. This metric exposes that the only material that can successfully attack all systems is Play-Doh, especially for the case of the optical sensor. The highest IAPMRs were obtained with silicone mixed with graphite, Play-Doh, and white glue.

For noncooperative attacks, the only vulnerable sensor was the thermal one for the case of silicone, latex, and Play-Doh, in very few times. It must be noted that very few attacks were performed in this manner, as it was decided to use the PCB molds that were created on the first try, with no room

for improvement, to see the results of a first-time attacker for this case.

In addition, the greater the APNRR is, the better the system is at rejecting fake samples (by not responding when they are placed on the sensor), so in this matter, the thermal sensor responds to the highest number of artefacts, although those captured samples ended up not being successful. Systems can also reject artefacts due to their low quality, and this ability is represented by APAR. In this case, the capacitive sensors were more capable of rejecting nonconductive samples, even when breathing on them to create a conductive layer on the surface.

5.2.2. Mobile Phones (Studies 2 and 3). In this case, only IAPMR was known, so the graphs are simplified by only showing this metric and omitting IAPNMR (its contrary).

First, the overall IAPMR results are shown in Figure 7. This first graph is shown inside the corresponding smartphone shape for clarification. The rest will be shown as usual for the sake of space. The figure compares the outcome obtained in study 2 (1 expert) with the average outcome obtained in study 3 (15 laymen attackers).

As it can be seen, the experienced attacker successfully attacked the smartphones more often than the inexperienced attackers, *in average*. The results for MD3 are quite similar. Nevertheless, the IAPMR for MD4 is higher for the inexperienced attackers than for the experienced one. This outcome will be explained later in this section.

The common materials for both studies were Play-Doh and gelatin, as they were tried on every smartphone by every attacker. Thus, Figure 8 shows a comparison of IAPMR results by an experienced attacker versus the average of 15

FIGURE 5: Cooperative attack results for desktop fingerprint sensors, separated by device and artefact species (material).

FIGURE 6: Noncooperative attack results for desktop fingerprint sensors, separated by device and artefact species (material).

unexperienced attackers. It can be clearly seen that the most vulnerable device is MD5, for both Play-Doh and gelatin. The experienced attacker had more successful attacks in most cases.

A break-down of the IAPMR results by attacker can be seen in Figure 9. The experienced attacker could hack into MD1 19% of the times, while the inexperienced had an almost

negligible number of successful attempts. MD2 was tricky to hack for all attackers, although Att 1 did a slightly better job. It can be observed that, for the case of MD3, IAPMR varies quite a bit depending on the inexperienced attacker, although the average is similar to the experienced one. On the case of MD4, the smartphone was hard to hack for most attackers (even for Att 1), but the inexperienced attacker Att 14 could break into

TABLE 9: Results on the vulnerabilities of each device to different artefact species. Not all materials were used in all sensors. A tick means that the device was successfully attacked with that species at least once. A cross indicates that the device could not be attacked with that material even once. An interrogation was used when the experiment was not tried.

Artefact species	Vulnerable devices								
	Ther	Cap1	Cap2	Opt	MD1	MD2	MD3	MD4	MD5
Play-Doh	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gelatin	✓	x	✓	✓	x	✓	✓	✓	✓
Clay	?	?	?	?	?	?	x	?	?
Wood glue	✓	x	x	x	?	?	✓	✓	?
Conductive ink	?	?	?	?	?	?	x	?	?
Latex	✓	x	x	x	?	?	✓	?	✓
Latex + graph	✓	x	x	✓	?	✓	✓	?	?
Silicone	✓	x	x	✓	?	?	x	?	✓
Silicone + graph	✓	x	x	✓	?	?	?	?	?

FIGURE 7: Overall IAPMR comparison per smartphone. Att 1 refers to the expert that performed the 2nd experiment, while Rest refers to the 15 simulated attackers.

it 19% of the times. Lastly, MD5 was the one with the most uneven results: Att 1 attacked the system successfully 40% of the times, while Att 15 did 25% of the times and Att 16, none.

The 15 simulated attackers had a chance to get additional credit for the assignment if they used additional artefact species and reported them. The results are shown in Figure 10, showing that it was discovered that white glue was the most successful material on MD3. Moreover, latex and silicone were found to be more successful than Play-Doh and gelatin on MD5.

In [21], it was said that attacking fingerprint sensors with fake fingers depends on expertise, but that luck also has a great impact. Sometimes, after trying to attack a sensor many times, the tester moves the finger slightly differently, or adds more water to the mix, or heats the artefact more and suddenly the fake finger works. This could also happen the first time a fake finger is used. Once the trick is known, the rest of the attempts will be much easier.

5.2.3. All Experiments. The only common metric for all experiments is IAPMR, which is the most significant one, as it shows the proportion of times that an artefact was verified as the real finger. It is shown for all devices across the 3 studies on Figure 11.

It can be observed that the highest IAPMR was obtained by the expert attacker on average on study 2, reaching a value of 40,2% for MD5. The lowest IAPMR on average happened on study 1 with the desktop sensors, as the policies for capture were stricter than those of the mobile devices. Our feeling when performing the evaluation on mobile devices after having done the same on desktop sensors was that it was much easier to bypass their security in comparison, which can be clearly noticed on the results.

5.3. Vulnerability Test Results. Although not all materials were tried on all devices, some vulnerabilities can be reported from the results obtained in the previous section (Table 9). A device being vulnerable means that it was hacked at least once with that artefact species. It can be observed that Play-Doh could successfully attack all devices at least once.

The consequences that may derive from these vulnerabilities are that an attacker could enter an unauthorized building, have access to privileges that do not belong to him/her, or for the case of the smartphones, unlock the phone, and have access to all the apps that do not require additional security.

5.4. Artefacts, Molds, and Captured Images. As it happens with cooking, every evaluator has different abilities to generate artefacts and it is very difficult to compare one's expertise to another. Therefore, the results of the security evaluation are dependent on who the attacker is. Thus, for a report, the best we can do is show which molds and artefacts were used in the evaluation.

5.4.1. Molds. Even inside the same evaluation, there are factors that condition the quality of the generated molds, and this influences highly the results. For example, some people tend to sweat a lot if their finger is surrounded by silicone paste for a few minutes, so the molds turn out with bubbles and with a very low quality (Figure 12).

5.4.2. Artefacts. As with the molds, it is important to include examples of the artefacts that were created of each artefact species. Figure 13 gives examples of the artefacts used for this experiment.

FIGURE 8: Comparison of IAPMR results by an experienced attacker (PLA1 for Play-Doh and GEL1 for gelatin) and 15 laymen attackers (PLA2 and GEL2).

FIGURE 9: Break-down of IAPMR results by each of the 15 unexperienced attackers, separated by attacker and device.

FIGURE 10: IAPMR results for the 15 inexperienced attackers, divided by all the artefact species used. PLA = Play-Doh; GEL = gelatin; CLA = clay; GLU = wood glue; LTX = latex; SIL = silicone.

FIGURE 11: Overall IAPMR for all devices across 3 experiments. Desktop sensors in this case only cover cooperative attacks, for the sake of fair comparison.

5.4.3. Captured Images. The only images that we could get access to were the ones from the desktop sensors evaluation. The examples on Figures 14 and 15 are from the thermal sensor.

The cooperative artefact has a notably better quality than the noncooperative one, due to its acquisition process. Cooperative molds capture quite accurately the shape of the capture subject's fingerprint, while getting an accurate sample from a latent print on a glass can be trickier, as many features can be lost on the process.

It must also be noted that even, within the same evaluation, the evaluator's ability can improve. For instance, making a good quality gelatin artefact consists of having the right proportion of water and gelatin leaves, and some trials might

need to be done before reaching that appropriate proportion. A clear example of this can be seen in Figure 15.

6. Conclusions

During the process of performing 3 separate PAD evaluations on fingerprint sensors, some lessons were learnt, the basic one being that standards and methodologies are necessary to compare PAD evaluations made by researchers and certification bodies. Studies are only comparable if the same metrics and procedures are used.

Enrolment and verification policies are very important for the performance and security of a system, and, with smartphones, these policies are out of hand for independent

FIGURE 12: Molds of different qualities. The mold on (a) has bubbles and is blurry due to the capture subject's finger characteristics.

FIGURE 13: Examples of 3 artefact species: latex with graphite, gelatin, and Play-Doh.

researchers and they must adapt to the ones given by the manufacturer. For instance, for the enrolment of mobile device MD5, only 6 captures are needed (while other devices use more than 15), and this poor enrolment policy could be a reason why it is noticeably more vulnerable to attacks than the others. Also, some devices allow an unlimited amount of attempts to present a sample, giving the attacker unlimited chances. This can be fixed by asking for additional information (PIN, password, or an additional biometric modality).

Attack potential is an adequate tool to measure the effort needed to attack a system. Nevertheless, in the case of

biometrics, it is difficult to calculate. Even within the same evaluation, the evaluator gets better at attacking the system in each attempt. Also, at any point of the process, anyone can get the trick to hacking a specific sensor (out of expertise and, mostly, out of luck) and the evaluation results can vary highly from that point on. To soften this variability, one solution is reporting examples of the molds, artefacts, and captured images of the evaluation.

Expertise was evaluated on this work. It was proved that all 5 smartphone sensors were hacked at some point by unexperienced attackers. The only prior knowledge they had was a short video of an expert creating fake fingers, so

FIGURE 14: Examples of captured images of cooperative and noncooperative artefacts from the desktop sensors evaluation.

FIGURE 15: Examples of a bad quality gelatin artefact versus a good quality gelatin artefact. Images from the desktop sensors evaluation.

this is especially preoccupying because similar videos can be obtained from the Internet easily. Fortunately, this was the case only for cooperative attacks, and their results were notably worse than the ones from the expert, in average. On the other hand, noncooperative attacks are more complex and need more expertise to be carried out. It was also noticed that it is easier to hack smartphone sensors than desktop sensors.

In the future, the database of inexperienced attackers will be increased, and further analyses will be made based on this. Moreover, more insights can be obtained in future evaluations: how long it takes for an inexperienced attacker to successfully attack a sensor for the first time, dependence

on finger used (index, middle, and thumb), dependence on left or right hand, and so on.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was partially supported by the Spanish National Cybersecurity Institute (INCIBE) under the Grants Program “Excellence of Advanced Cybersecurity Research Teams.”

Supplementary Materials

The supplementary video is the same file that was shared with the inexperienced attackers in order to perform their experiments. It shows the whole process from creating a mold from a real finger, to creating an artefact and using it on the smartphone sensor. (*Supplementary Materials*)

References

- [1] R. Blanco-Gonzalo, R. Sanchez-Reillo, J. Liu-Jimenez, and C. Sanchez-Redondo, "How to assess user interaction effects in biometric performance," in *Proceedings of the IEEE International Conference on Identity, Security and Behavior Analysis (ISBA '17)*, February 2017.
- [2] D. He and D. Wang, "Robust biometrics-based authentication scheme for multiserver environment," *IEEE Systems Journal*, vol. 9, no. 3, pp. 816–823, 2015.
- [3] L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2021–2040, 2003.
- [4] T. van der Putte and J. Keuning, "Biometrical fingerprint recognition: do not get your fingers burned," in *Proceedings of 4th Working Conference on Smart Card Research and Advanced Applications, IFIP TC8/WG8*, vol. 31, p. 16, Bristol, UK, September 2000.
- [5] S. A. C. Schuckers, "Spoofing and anti-spoofing measures," *Information Security Technical Report*, vol. 7, no. 4, pp. 56–62, 2002.
- [6] A. Choiniere and T. Lubysheva, Novetta—Protecting against Fingerprint Vulnerabilities when Deploying Biometric Systems, 2015.
- [7] J. Blommé, "Evaluation of biometric security systems against artificial fingers," 2003.
- [8] L. Ghiani, D. A. Yambay, V. Mura, G. L. Marcialis, F. Roli, and S. A. Schuckers, "Review of the Fingerprint Liveness Detection (LivDet) competition series: 2009 to 2015," *Image and Vision Computing*, vol. 58, pp. 110–128, 2017.
- [9] A. Wiehe, T. Søndrol, O. K. Olsen, and F. Skarderud, "Attacking fingerprint sensors," 2004.
- [10] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial 'gummy' fingers on fingerprint systems," in *Proceedings of the Optical Security and Counterfeit Deterrence Techniques IV*, vol. 4677 of *Proceedings of SPIE*, pp. 275–289, January 2002.
- [11] J. Galbally-Herrero, J. Fierrez-Aguilar, J. D. Rodriguez-Gonzalez, F. Alonso-Fernandez, J. Ortega-Garcia, and M. Tapiador, "On the vulnerability of fingerprint verification systems to fake fingerprints attacks," in *Proceedings of the 40th Annual IEEE International Carnahan Conference on Security Technology*, pp. 130–136, October 2006.
- [12] S. Schuckers, "Presentations and attacks, and spoofs, oh my," *Image and Vision Computing*, vol. 55, pp. 26–30, 2016.
- [13] Frank, "Chaos Computer Club breaks Apple TouchID," 2013, <https://www.ccc.de/en/updates/2013/ccc-breaks-apple-touchid>.
- [14] A. Hern, "Hacker fakes German minister's fingerprints using photos of her hands," *The Guardian*, 2014.
- [15] Common Criteria, "Common Criteria for information technology security evaluation part 1: introduction and general model September 2012 revision 4," ISO/IEC 15408 Common Criteria, Part 12012, International Organization for Standardization and International Electrotechnical Commission, 2012.
- [16] Common Criteria, "Common methodology for information technology security evaluation methodology September 2012 revision 4 foreword," Tech. Rep. Ccmb-2012-09-004, 2012.
- [17] ISO/IEC/JTC 1/SC37, *Text of FDIS 30107-3, Information Technology—Biometric Presentation Attack Detection—Part 3: Testing And Reporting*, vol. 2008, ISO-IEC Standards, 2009.
- [18] B. Fernandez-Saavedra, R. Sanchez-Reillo, J. Liu-Jimenez, and J. G. Ruiz, "Best practices for the security evaluation of biometric systems," in *Proceedings of the 48th Annual IEEE International Carnahan Conference on Security Technology*, October 2014.
- [19] B. Fernandez-Saavedra, R. Sanchez-Reillo, R. Alonso-Moreno, and C. Sanchez-Avila, "Evaluation methodology for fake samples detection in biometrics," in *Proceedings of the 42nd Annual 2008 IEEE International Carnahan Conference on Security Technology*, pp. 233–240, October 2008.
- [20] B. Fernandez-Saavedra, R. Sanchez-Reillo, C. Sanchez-Redondo, and R. Blanco-Gonzalo, "Testing of biometric systems integrated in mobile devices," in *Proceedings of the 49th Annual IEEE International Carnahan Conference on Security Technology*, pp. 321–326, September 2015.
- [21] I. Goicoechea-Telleria, J. Liu-Jimenez, R. Sanchez-Reillo, and W. Ponce-Hernandez, "Vulnerabilities of biometric systems integrated in mobile devices: an evaluation," in *Proceedings of the 50th Annual IEEE International Carnahan Conference on Security Technology*, October 2016.
- [22] I. Goicoechea-Telleria, B. Fernandez-Saavedra, and R. Sanchez-Reillo, "An evaluation of presentation attack detection of fingerprint biometric systems applying ISO/IEC 30107-3," in *Proceedings of the International Biometric Performance Testing Conference*, 2016.
- [23] T. Dunstone and N. Yager, *Biometric System and Data Analysis: Design, Evaluation, and Data Mining*, Springer, 2009.
- [24] Y. Jo, S. Jeon, J. Im, and M. Lee, "Security analysis and improvement of fingerprint authentication for smartphones," *Mobile Information Systems*, vol. 2016, Article ID 8973828, 11 pages, 2016.
- [25] E. Pagnin and A. Mitrokovska, "Privacy-preserving biometric authentication: challenges and directions," *Security and Communication Networks*, vol. 2017, Article ID 7129505, 9 pages, 2017.
- [26] C.-I. Fan and Y.-H. Lin, "Provably secure remote truly three-factor authentication scheme with privacy protection on biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 933–945, 2009.
- [27] S. Yoon, J. Feng, and A. K. Jain, "Altered fingerprints: analysis and detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 3, pp. 451–464, 2012.
- [28] A. K. Jain and S. Yoon, "Automatic detection of altered fingerprints," *Computer*, vol. 45, no. 1, pp. 79–82, 2012.
- [29] S. Ghouzali, M. Lafkih, W. Abdul, M. Mikram, M. El Haziti, and D. Aboutajdine, "Trace attack against biometric mobile applications," *Mobile Information Systems*, vol. 2016, Article ID 2065948, 15 pages, 2016.
- [30] L. Thalheim, J. Krissler, and P.-M. Ziegler, "Body Check: Biometrics Defeated," 2002, <http://www.pcmag.com/article2/0,2817,13919,00.asp>.
- [31] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers, LivDet 2009—Fingerprint Liveness Detection Competition 2009, 2009.

- [32] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers, "LivDet 2011—fingerprint liveness detection competition 2011," in *Proceedings of the 2012 5th IAPR International Conference on Biometrics (ICB '12)*, pp. 208–215, April 2012.
- [33] L. Ghiani, D. Yambay, V. Mura et al., "LivDet 2013 fingerprint liveness detection competition 2013," in *Proceedings of the 6th IAPR International Conference on Biometrics (ICB '13)*, pp. 1–6, Madrid, Spain, June 2013.
- [34] V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay, and S. A. Schuckers, "LivDet 2015 fingerprint liveness detection competition 2015," in *Proceedings of the 7th IEEE International Conference on Biometrics Theory, Applications and Systems (BTAS '15)*, September 2015.
- [35] K. Cao and A. K. Jain, "Hacking Mobile Phones Using 2D Printed Fingerprints," 2016.
- [36] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Computing Surveys*, vol. 47, no. 2, article 28, 2014.
- [37] D. Zhang, Z. Guo, and Y. Gong, *Multispectral Biometrics: Systems and Applications*, Springer, 2015.
- [38] NIST, "NIST Fingerprint Image Quality," *Biometric Quality*, <https://www.nist.gov/programs-projects/biometric-quality-homepage>.
- [39] ISO/IEC/JTC 1/SC37, "DRAFT INTERNATIONAL STANDARD ISO/IEC DIS 30107-3 Information technology—Biometric presentation attack detection," 2017.
- [40] NIST, NIST Biometric Image Software, 2015, <http://www.nist.gov/itl/iad/ig/nbis.cfm>.