Security Impact of High Resolution Smartphone Cameras

Tobias Fiebig[*] and Jan Krissler[*]
Technische Universität Berlin
FG Security in Telecommunications
Berlin, Germany
{*tfiebig,starbug*}*@sec.t-labs.tu-berlin.de*

Ronny Hänsch[*]
Technische Universität Berlin
FG Computervision
Berlin, Germany
r.haensch@tu-berlin.de

Abstract

Nearly every modern mobile device includes two cameras. With advances in technology the resolution of these sensors has constantly increased. While this development provides great convenience for users, for example with video-telephony or as dedicated camera replacement, the security implications of including high resolution cameras on such devices has yet to be considered in greater detail. With this paper we demonstrate that an attacker may abuse the cameras in modern smartphones to extract valuable information from a victim. First, we consider exploiting a front-facing camera to capture a user's keystrokes. By observing facial reflections, it is possible to capture user input with the camera. Subsequently, individual keystrokes can be extracted from the images acquired with the camera. Furthermore, we demonstrate that these cameras can be used by an attacker to extract and forge the fingerprints of a victim. This enables an attacker to perform a wide range of malicious actions, including authentication bypass on modern biometric systems and falsely implicating a person by planting fingerprints in a crime scene. Finally, we introduce several mitigation strategies for the identified threats.

1 Introduction

In recent years, smartphones have become ubiquitous and their popularity continues to grow. While sales of PC hardware continue to decline, sales of mobile hardware continue to increase. More and more applications, like banking and mobile payment services, now target mobile platforms as well. Social networks an messaging services are also extremely popular on mobile devices. As a result, modern smartphones contain significant amounts of sensitive data.

With ever-increasing features, new sensors and peripherals are continuously integrated into these systems, the most notable of which are the multi megapixel front- and rear-facing cameras. However, attackers lack the capability of accessing these peripherals directly. Modern mobile operating systems implement fine-grained permission systems to prevent unauthorized access. Users can choose to permit or deny access to certain peripherals at the time of the installation. However, such access restrictions are ineffective as many users will agree to grant access to malicious applications that they willingly install on their devices.

Work by Felt et al. published in 2011 has placed the camera permission in the top 10 of unnecessary, yet commonly requested permissions across the Android app market [7]. Other publications have concluded "that the majority of Android users do not pay attention to or understand permission warnings" [8]. Based on this information a malicious attacker can gain remote access to the device's cameras when a user installs a malicious application. Moreover, numerous local root exploits exist for popular Android devices, allowing an attacker to gain system privileges. This allows an attacker to completely bypass user permission requests [29].

Even though an attacker can gain access to a camera, they are seldom considered security-relevant sensors. The primary focus of previous research has been on the associated privacy issues [9, 4]. In particular, Simon and Anderson investigated privacy issues of the front facing camera in smartphones by developing an orientation based keylogger [24]. However, the resolution of smartphone cameras make attacks more and more feasible with each generation (see Section 3). Hence, we chose to use the highest resolution cameras currently available on the market. This allows us to evaluate the current as well as future effectiveness of the presented attacks.

[*] These authors contributed equally in their respective fields.

The main contributions of our work are the following:

Facial Reflection Keylogger. We present a new method utilizing facial reflections for recording a user's keystrokes. By continuing from work of Xu et al. [27], this technique surpasses the performance of previous front-camera based keyloggers [24] and removes the need of physical proximity [27]. Our evaluation on a real device furthermore proves that a complex framework is not necessary for a real-world attack as the extraction process can be performed manually. This allows an attacker to circumvent even most advanced anti malware and keylogging mechanisms like separate operating system compartments on the most recent high-security smartphones.

Fingerprint Extraction. We present a new method to extract a user's fingerprints with a mobile phone's camera for creating forgeries. With these forgeries used in our experiments, we were able to bypass the most recent fingerprint readers found in smartphones, as well as traditional fingerprint sensors. An attacker could use these prints to gain access to biometrically secured areas or implicate a victim in a crime by placing false prints on a crime scene. Furthermore, these techniques enable an attacker to circumvent modern fingerprint-based authentication methods, such as those which now become common on modern smartphones [25]. While traditional methods rely on a fingerprint being found on the device itself [5], an attacker can prepare a forgery before obtaining the phone. This solves the issues that a well preserved fingerprint may be unavailable on the device surface and relevant data may be remotely wiped if the user considers the phone misplaced or stolen.

Structure

The remainder of this paper is structured in the following manner. We first introduce the necessary background for our attacks in Section 2. This section also contains our description of a possible attack model. Section 3 describes the facial reflection attack vector for keylogging via the front-camera of modern smartphones. In the same way, Section 4 presents our work on the extraction and replication of fingerprints with a mobile phone's rear-facing camera. We will continue by discussing the results of both sections in Section 5. This section also contains the possible mitigation strategies, to prevent the exploitation of the uncovered issues. Section 6 then holds our final conclusion based on the evaluation in Section 5.

2 Background and Related Work

Due to their unique characteristics, different threats apply to the front- and rear-facing cameras on smartphones. To fully cover both types of cameras, we introduce a reflection based keylogger that abuses the front-camera and a technique to extract fingerprints with the rear-facing camera. Hence, we present the required background and related work for our reflection based keylogger first. We then continue by introducing the background necessary to understand the fingerprint extraction technique using the back-facing camera of a smartphone. To provide some practical context for our work we then introduce an attacker model for the issues we discovered.

2.1 Visual Keylogging

Keyloggers are small programs installed on a computer system that extract input information. Due to the unique nature of mobile devices, the acquisition of input information from mobile devices poses a greater challenge to attackers than on the PC platform. On one hand, most of these devices lack a physical input method, i.e. keyboard. Instead, a soft-keyboard is used in these systems, which is displayed on a touch screen and records input based on the section of the screen being touched. On the other hand, advanced privilege separation models have been implemented on mobile platforms, which restrict an attacker from accessing security sensitive functions.

Nevertheless, various publications in recent years have shown that such attacks are possible. While the initial publications are mostly concerned with obtaining touch-input information from the touch-screen of a mobile device [23, 6] these attack vectors have been mitigated on modern devices. New attack vectors applying heuristic approaches on sensory information available on mobile devices were subsequently discovered. Most prominently, the accelerometer of phones was exploited to determine the orientation of a mobile device [1, 28].

Visual keylogging utilizing cameras is based on two important strategies. The first strategy has been introduced by Simon and Anderson [24] in 2013. With their work, they have demonstrated the privacy issues that arise from the presence of front facing user cameras as orientation information can be reconstructed from pictures taken with it. Following the concept of various accelerometer based keyloggers presented earlier [1, 28], the user's keystrokes can be extracted from the images. Although their work is promising, the limited orientation resolution extracted from images leads to serious limitations. This means that their approach is limited to $3 \times 3 + 1$ numerical keypads and they could not identify a key press with significantly more than 50% certainty in their empirical study [24].

The second strategy is the visual eavesdropping on computer screens and smartphones, also known as shoulder surfing. This technique is concerned with capturing the screen from a relative distance while the screen content and possible inputs can still be reconstructed. Kuhn et al. published one of the earliest works on this matter in 2003 [17], where the authors reconstruct a $32 \times 24cm$ display from a distance of $60m$ by using a professional telescope. Later works (e.g. Backes et al. [2]) extended existing approaches by utilizing reflections and consequently overcame the requirement of a direct line-of-sight between observer and target. While the proposed methods were able to successfully recover the typed input in the case of direct line-of-sight attacks, the reconstruction accuracy decreased significantly in cases of even a single reflection. In 2013, Xu et al. [27] proposed a new approach that neither depends on the detection of small visual details, nor on a direct line-of-sight. Instead of trying to reconstruct the whole screen content, the method tracks the user's fingers as they move over the screen. The relation of the movement, i.e. pauses, and the position of the fingertips is used afterwards to reconstruct the typed input.

For this purpose, Xu et al. created a semi-automatic framework which analyzes a video file in a multi-step process, automatically performing the steps mentioned in the previous paragraph [27]. They carried out a range of experiments with varying settings, evaluating distance between $30 - 50m$ in direct line-of-sight, to $4 - 10m$ distances fore single and $3m$ distance double reflection scenarios. The general results are very promising with 23% (17/73) of the test sentences being perfectly reconstructed, while 92% of them have an METEOR score ([18]) above 0.5, which means that they are still understandable by a human. Of the 15 selected test passwords, 12 have been reconstructed in 12 or fewer guesses and no password needed more than 6000 guesses. For our comparisons we choose their experiments with a Canon VIXIA camcorder as documented in Table 1. The results of their experiments with this camera in a single reflection on sunglasses setting state a perfect reconstruction for a distance of $4m$ and a well understandable reconstruction (METEOR-score of 0.71) for one of $10m$. The distance between object and reflecting surface is assumed to be $30cm$ in both cases.

As can be seen, the two most important factors of the approach are the quality of the camera and the distance between observer and target. Taking these two factors into account, the issue boils down to the resulting size of the target device in the recording. Xu et al. [27] created two formulas to calculate the target's size in the resulting image based on these factors. The target size in pixels per axis in the captured image for a direct observation can be calculated by Equation 1 [27].

$$Size_{Direct} = \frac{SensorResolution}{SensorSize} \cdot \frac{ObjectSize}{\frac{TargetDistance}{FocalLength} - 1} \quad (1)$$

If, however, reflections are involved, the curvature of the reflective surface and the distance of the reflective surface from the target have to be taken into account as well. This leads to Equation 2 [27].

$$Size_{Reflection} = Size_{Direct} \cdot \frac{1}{\frac{2 \cdot DistanceFromSurface}{CurvatureRadius} + 1} \quad (2)$$

2.2 Fingerprints and Biometrics

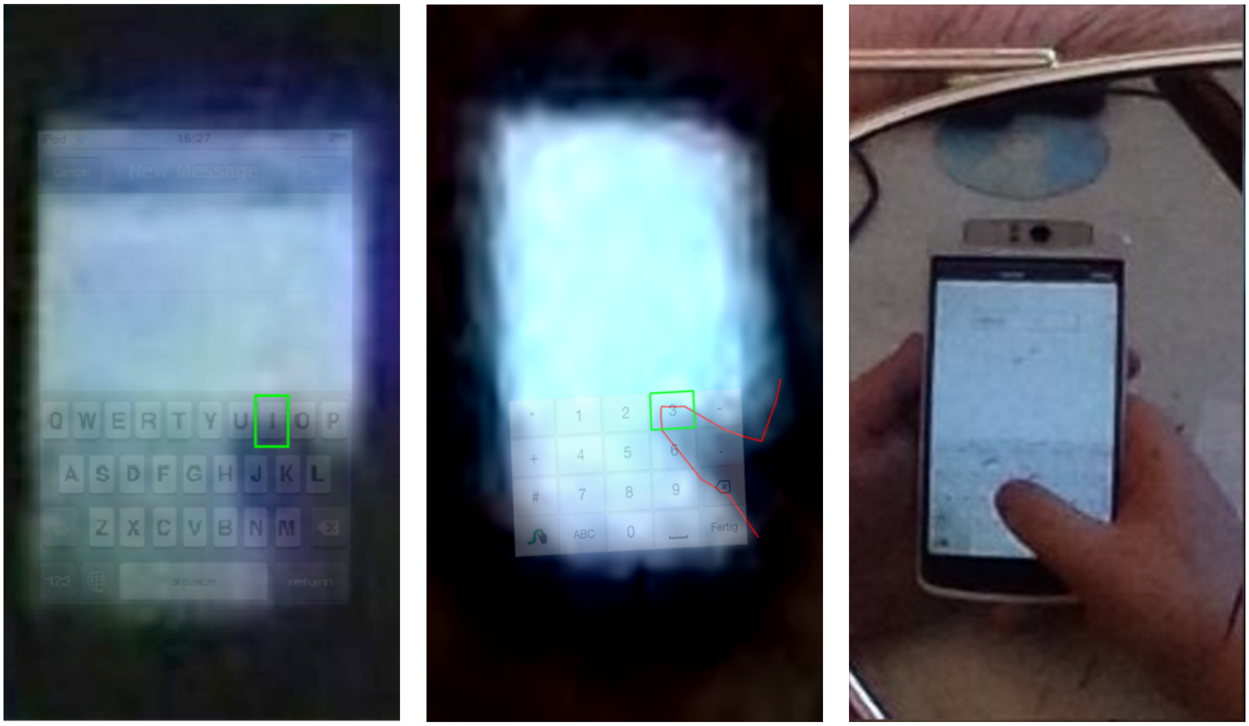
Fingerprints are the oldest biometric feature that has been actively used [16]. Fingerprints on ancient seals and clay tablets are even actively used in the field of archaeology [14]. In modern times, they are used in forensics to identify a perpetrator among the of suspects of a crime [19].

The question how fingerprints do form during the pregnancy has not yet been conclusively solved [16, 15]. What however is certain is that folds of the epidermal layer leads to so called ridges and valleys on the outside of fingers and feet. Until a couple of years ago, fingerprints were taken using ink and paper and were compared manually based on global structures like whirls, arches or loops [19]. These patterns are formed by the ridges and valleys of a fingerprint.

Their widespread use started with the development of digital sensors and associated image recognition libraries. These sensors are based on a variety of techniques for capturing the fingerprints, but only two have prevailed [22]. Optical sensors use the physical principle of scattered total reflection [10]. They have a high resolution, but are relatively large and are therefor mostly found at static stations like for example border controls or access to buildings.

Capacitive sensors on the other hand are considerably smaller. They rely on measuring the difference in the capacity between the skin where the the ridges directly touch the sensor and the air between sensor and skin in the valleys [26]. Due to their smaller size, they are mainly found in mobile devices like notebooks, and nowadays smartphones.

New sensors also use an additional RF (Radio Frequency) field to measure deeper skin layers. When the finger touches the sensor, an electrical field penetrates the finger and is reflected on lower layers of the skin. Hence the sensor images not the surface, but lower skin layers and is therefore more resilient against dirt and injuries of the upper skin layer [22].



(a) Canon VIXIA, Sunglasses with overlay from Xu et al. [27]

(b) OPPO N1, Eyeball with overlay

(c) OPPO N1, Sunglasses no overlay

Figure 1: Comparison of different images used in keystroke recovery. The quality of the reflection for the case of the Canon VIXIA (a) recording a reflection on sunglasses in 10m distance as obtained by Xu et al. [27] produces reflections comparable to the case of the OPPO N1 retrieving reflections from a user's eyes (b). Without applying a framework, it can be determined that the user in (b) presses a 3 on a numerical keypad. Figure (c) demonstrates the quality of reflections on sunglasses obtained with the OPPO N1. Please note that the used QWERTY keyboard can be clearly identified as such.

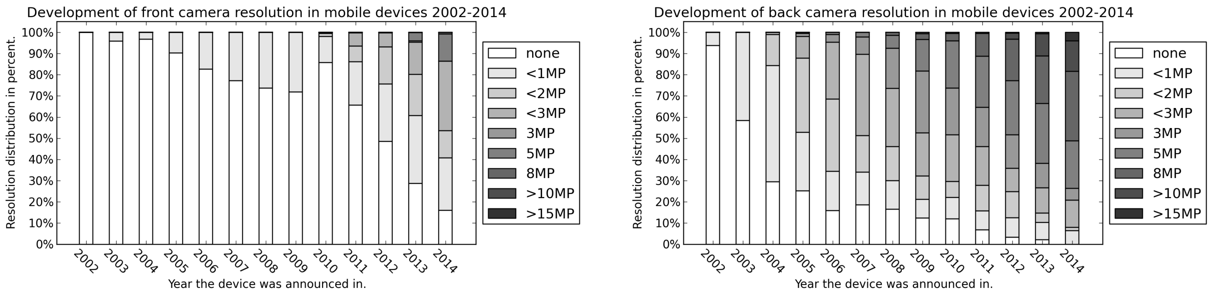
2.3 Attack Model

As for this attack model, let us consider one of the most high profile targets using the most advanced security mechanisms, being challenged by an equally advanced attacker. This means that we will investigate how the secretary of defense might be targeted by a foreign agency with the goal of stealing state secrets of outmost importance. Being aware of the constant threat of espionage, the ministry decided to issue high security phones with separate compartments for private and confidential use[1], secured by a pin-code and a fingerprint of the user. Within this high profile situation, the secretary of defense also decides to keep confidential documents in a fingerprint and combination secured safe in the office, instead of relying on a physical key, which might get stolen. As the combination is rather complex, he however notes it down in the confidential compartment of his smartphone.

By publishing a rather sophisticated malware posing as a harmless game, but sneaking in the permission to use the systems camera, they can infect a vast amount of phones, including the compartment for private use on the victims device. While they can use their foothold application to obtain root access on the private compartment, the confidential compartment and the pin-pad for unlocking it remain out of reach. While the rear-facing-camera is used to extract the targets fingerprints, further information from the private sector indicates that the target will go on an rather sunny holiday trip. This gives the attacker the opportunity to use a facial reflection based keylogger to extract the pin-code entered on the secure compartment while the target wears sunglasses.

In a final sweep, the foreign agency obtains all the confidential data they desire. While the target is on vacation, a targeted thief retrieves the victims phone. As recommended, the phone is encrypted, but with a replica of the extracted fingerprint and the previously extracted pin-code, the attackers can quickly recover all confidential data. When the target notices the theft 15 minutes

[1] e.g.: `www.viasat.com/mobile-enterprise-security`, `www.blackberry.com/secureworkspace`, `www.samsung.com/knox-mobile`



(a) Front camera resolution development.

(b) Rear-facing camera resolution development.

Figure 2: Development of the resolutions for front- (a) and rear-facing (b) cameras in announced devices between 2002 and 2014. One can see that the development for front-facing cameras is roughly six years behind. Hence, front-facing cameras with 16MP and more can be expected for 2018-2020. The graph has been generated with data gathered from gsmarena.com end of Feburary 2014.

later, and immediately issues a remote wipe, the phone is successfully wiped. However, the data, including the combination for the safe in the office, has already been stolen. Shortly after this incident, bribed cleaning personnel uses the recovered combination and a fingerprint replica created from the prints extracted with the mobile device to empty the safe in the office. As it turns out, the attackers were able to extract all confidential data from the phone and conveniently got access to the documents stored in the safe in the target's office. To hide their actions behind confusion, the attackers then use the forged fingerprints on a knife that is used in a murder. With the secretary of defense implicated in a crime, the whole incident goes unnoticed within the ensuing scandal.

3 Front-Camera Based Visual Keylogger

In this section, we demonstrate how techniques on reflection based shoulder surfing can be combined with a smartphone's front camera to construct a reflection based keylogger. To that end, we first consider algorithms proposed in related works. Based on these metrics, we are able to to demonstrate that our approach outperforms other solution. Furthermore we describe a set of experiments demonstrating that such a framework is not even necessary as the input can be obtained manually.

3.1 Theoretical Applicability

Utilizing the equations presented in Section 2, we can calculate if and with which expected accuracy, the framework as proposed by Xu et al. [27] is applicable to the new attack vector which we have identified. We decided to utilize the case of the Canon VIXIA in a 4m distance one-time reflection scenario for this comparison. For both cases, we assume an OPPO N1 with a $13cm \times 7.5cm$ screen to be the target device. Equations 3 and 4 present the estimated (x,y)-target size in the source-image for that case using Equation 2.

$$X_{VIXIA} = \frac{1920px}{4.84mm} \cdot \frac{130mm}{\frac{4000mm}{57mm}-1} \cdot \frac{1}{\frac{2\cdot 300mm}{8mm}+1} \cong 9.80 \tag{3}$$

$$Y_{VIXIA} = \frac{1080px}{3.42mm} \cdot \frac{75mm}{\frac{4000mm}{57mm}-1} \cdot \frac{1}{\frac{2\cdot 300mm}{8mm}+1} \cong 4.50 \tag{4}$$

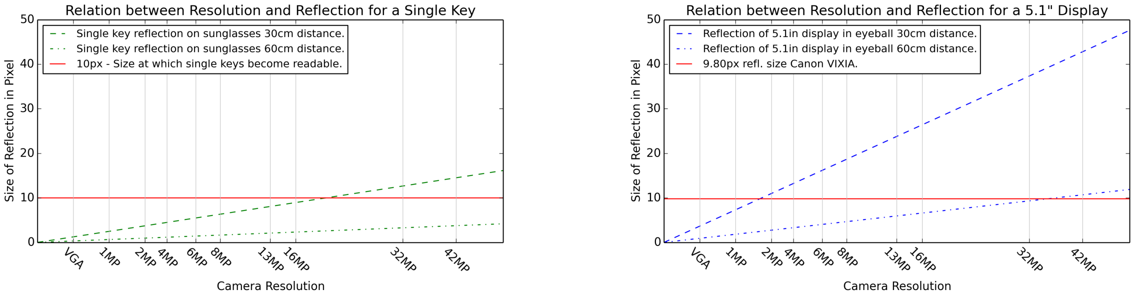
In this case the resulting image would have a $Size_{Reflection}$ of approximately $9.80px \times 4.50px$. If we now consider the case of the OPPO N1 in a scenario where the phone is used on an reflection of itself in the user's eye, we get Equations 5 and 6.

$$X_{OPPO} = \frac{4160px}{4.4mm} \cdot \frac{130mm}{\frac{300mm}{5mm}-1} \cdot \frac{1}{\frac{2\cdot 300mm}{8mm}+1} \cong 27.41 \tag{5}$$

$$Y_{OPPO} = \frac{3120px}{3.6mm} \cdot \frac{75mm}{\frac{300mm}{5mm}-1} \cdot \frac{1}{\frac{2\cdot 300mm}{8mm}+1} \cong 14.50 \tag{6}$$

As can be seen, $Size_{Reflection}$ is with $27.41px \times 14.50px$ nearly nine times larger in our case. Using less curvy reflection surfaces like sunglasses would even lead to a larger value for $CurvatureRadius$, hence even larger values for $Size_{Reflection}$. A comparison of the observable quality for the presented sensors and reflective surfaces is depicted in Figure 1. Hence, it is save to assume that results obtained with this technique will reach at least the accuracy observed by Xu et al. [27], which was already close to the ideal case of obtaining all inputs without any error.

Although the OPPO N1 has a 13MP camera, which can be front-facing, most devices on the market do not yet have front-facing cameras with a resolution that high.



(a) Reflection of single key for $30-60cm$ distance on sunglasses.

(b) Reflection of a 5.1” device for $30-60cm$ distance in an eye.

Figure 3: Size of an objects in the recorded image based on the distance between the phone and the face as well as the utilized reflection. Figure (a) depicts the size of a single key in the recorded image when reflections on sunglasses are used. The red line indicates when different keys become distinguishable. Cameras with a resolution between 16MP and 32MP suffice to actually read the screen content from a reflection. In Figure (b) the size of an reflection of a 5.1” display in the user’s eyes is depicted. The red line indicates the reflection size used by Xu et al. [27] for nearly perfect reconstruction with their framework. This demonstrates that cameras with only 2MP are already sufficient for corneal keylogging if the phone is held in not more than $30cm$ distance. Cameras of 32MP even allow for keylogging operations if the phone is held at $60cm$ distance.

As depicted in Figure 2(a), the most common resolutions are around 2MP. Using Equation 2, we investigated with which resolutions an corneal reflection can be used for effective keylogging. Using the case of the Canon VIXIA as the base line, we can prove that for a 5.1” device held at 30cm distance a corneal keylogger is feasible. With increasing resolutions, the distance between the device and the face can be increased as visualized in Figure 3(b). Assuming that a size of 10px in the recorded image is sufficient to distinguish different keys, we determined that sunglasses in conjunction with future higher resolutions enable an attacker to actually read the used keyboard. As depicted in Figure 3(a), we determined that a camera with 16MP to 32MP would be sufficient to actually read the keys on the keyboard. Following the graphs in Figure 2, such resolutions for front-cameras can be expected to enter the market starting 2018.

Feature	Canon VIXIA	OPPO N1 [21]
Resolution	$1920 \times 1080px$	$4160 \times 3120px$
SensorSize	$1/3.2$“	$1/3.06$“
	$= 4.84 \times 3.42mm$	$= 4.4 \times 3.6mm$
Focal Length	$57mm$	$5mm$
Target Dist.	$4 \cdot 1000mm$	$300mm$
Surface Dist.	$300mm$	$300mm$
Object Size	$73 \times 130mm$	$73 \times 130mm$

Table 1: Base characteristics of the Canon VIXIA used by Xu et al. [27] and the OPPO N1 camera. The lower focal length of the OPPO N1 is more than compensated by the significantly larger resolution of the sensor.

3.2 Experimental Verification

We conducted a set of experiments with the OPPO N1 and its 13MP camera (as described in Table 1) to demonstrate that a manual extraction is as feasible as the automated process already documented in the literature.

For this experiments, we created an application which provides a user with a numerical password input field. On each keypress, the application records an image with the front camera in the background. These images were stored on the phone for later extraction. A second subject was then tasked with determining the entered pins based on the provided images, using only the reflections found in the user’s eyes. The results of this process can be found in Table 2. While the correct pincode was easily established for two out of four cases, the two other cases demonstrate that a permutation over the most probable as well as second most probable input may be necessary to establish the correct pin.

Entered	First Guess	Second Guess
78135	48435	78128
90134	60121	99254
5102	5102	/
159397	159397	/

Table 2: Results of the manual pin-code recovery. Two were recognized on the first try while a second round of guesses was necessary for the other two.

3.3 Previous Mitigation Techniques

In the context of facial reflection shoulder surfing, some mitigation strategies were suggested by Xu et al. [27]. However, all three techniques proposed by them do not sufficiently prevent the attacks we just described. The first technique they propose is a privacy screen limiting the view-port of a device, the second one are gaze-based passwords and the third are randomized keyboards. The privacy screen does not provide additional protection against our method, as the reflections are created in the direct view port, i.e. the face of the user. Gaze based passwords would be entered via eye-tracking with the front camera, the sensor which we already utilize for our attack.

Finally, randomized keyboards do provide some protection, but only as long as the recorded image does not have a resolution high enough to actually read the random keyboard. However, according to our analysis in the previous sub-section, even slight increases in the resolution of user-facing cameras in conjunction with worn sunglasses will provide sufficient images. Based on the data presented in Figure 2(a) and (b), such devices can be expected in the near future. For lower resolution cameras, it might be feasible to measure the time a user needs to press a key on the randomized keyboard. If that time is closer to the subjects native typing speed, it may be an indication that the letter for that key is on its native keyboard position. Over time, it would hence be possible to extract a password, each character when the corresponding key is in its native position.

4 Rear-Facing Camera Fingerprint Extraction

In this section we introduce a new method to extract a user's fingerprints and demonstrate how these extractions can be used to create forgeries sufficient to break the most recent mobile fingerprint readers. During these experiments we used the OPPO N1 as introduced in Section 3. As we exploit the camera of a mobile device for this, the whole process requires no physical contact to the victim. Our successful creation of forgeries demonstrates that cameras in smartphones are a threat to biometric authentication mechanisms.

4.1 Fingerprint Extraction

The first to effectively clone a fingerprint usable on a fingerprint sensor was Matsumoto in 2002 [20], although the forensic literature holds indications of forgeries being conducted well before that [11]. Matsumoto used gummy and rubber replicas to create forgeries from either mold-prints of real fingers or laser printed negatives of scanned fingerprints produced by pressing an inked finger on a sheet of paper. We discovered that the resolution of modern smartphone cameras suffices to obtain images of fingers that can be used for the same process. In an attack scenario an adversary has to obtain an image of the main-hand (either left or right) index finger, the one mostly used for biometric authentication. We achieved the best results when the phone was put down with the front side facing the table. During the subsequent pick-up by the user, ideal images of the main hand index finger may be created. This process is depicted in Figure 4.

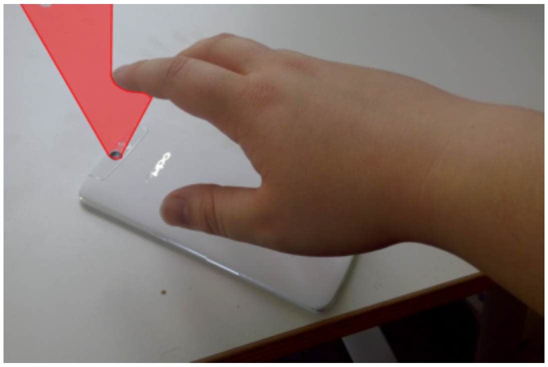
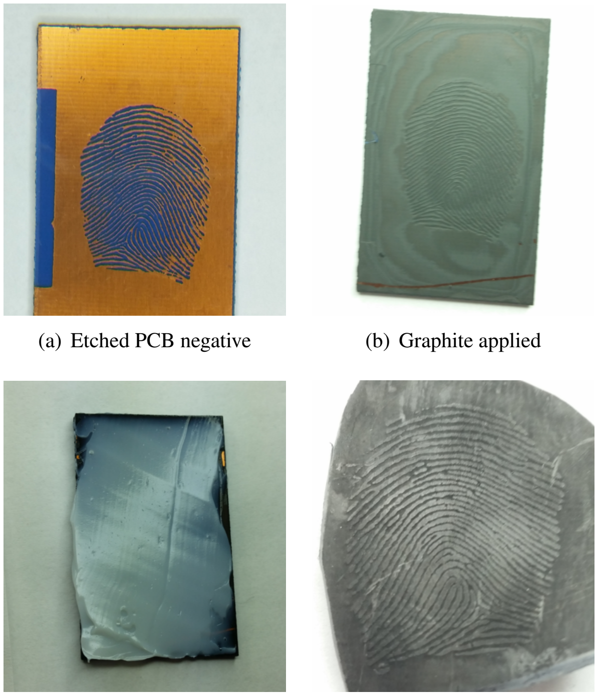


Figure 4: While a user picks up a device the right index finger moves through the rear-facing camera view-port (red). An attacker can use this moment to create an image of the user's fingerprints.

4.2 Cloning Process

The images taken with the technique described above are the basis for the cloning. As depicted in Figure 6(a) and 6(b) this process consists of first identifying the section of the image which contains the fingerprint to extract. Then the image is manipulated by a binary filter, which transforms the darker valleys to black and the lighter ridges to white. Furthermore all peripheral parts of the image are cropped and replaced by a black area surrounding the print. As this does not necessarily yield perfect results, manual intervention can be required to adjust fine-grained parts of the print.

In contrast to direct scans of latent prints, the real size of an object taken by a camera is not known. It depends on the zoom and distance between finger and camera. To get an estimation of the finger's size, additional information like zoom level or the auto-focus settings could be used. As biometric features change over time and even between single measurements subtle changes occur, most systems allow some tolerance. The fingerprint system from Digital Persona we used for testing tolerated a size variation of $+-10\%$ during our empirical evaluation.



(a) Etched PCB negative

(b) Graphite applied

(c) Wood-glue applied

(d) Ready forgery

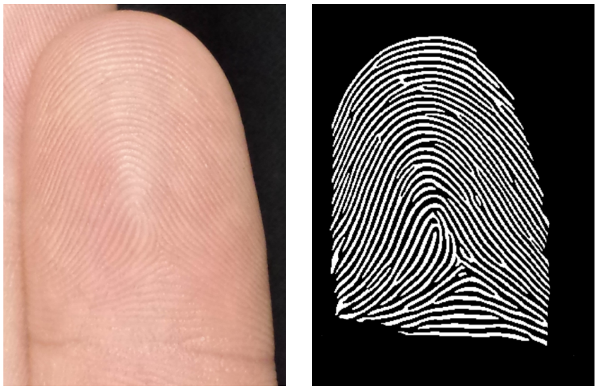
Figure 5: The four stages during forgery. (a) First a negative is etched from a PCB. (b) Then graphite spray is applied to allow for easier peeling of the forgery, and to adjust the capacity of the wood-glue. (c) The applied wood glue on the negative as to set, this usually takes around one hour. (d) The created forgery can then be used on the designated target. ready for use.

To create a mold for the dummies, we first enhanced the pictures in contrast and brightness until the ridges and valleys are distinguishable for the binarisation step. By splitting the values of the brightness channel valleys turn black and ridges white as depicted in Figure 6(a) and 6(b). Depending on the quality of the image some manual post-processing has to be performed. As a picture of the finger is taken, the resulting images must be mirrored before they can be printed onto a transparent foil using a laser printer. Modern sensors have a resolution of at least 500dpi, but up to 1000dpi [12], so the print-out should have at least this resolution. The toner particles form a three dimensional structure with a height of around 15 microns, which is sufficient to fool most types of sensor.

Thermal sensors use the different cooling-time between the air in the valleys and the skin of the ridges to create an image of a fingerprint. Modern capacitive sensors emit an additional RF field into the finger, which allows them to measure deeper skin layers. To account for these features the dummies for such sensors have to be created with deeper molding structures. To create these we used the print-out as an etching-mask on a photo sensitive printed circuit board (PCB). These boards come with a copper layer of 35 or 70 microns which is approximately the height of the ridges in a human fingerprint. Subsequently, the PCB is etched to remove the undesired areas. Finally, the resulting dummy can be used to create replicas by applying a thin layer of common wood-glue on it. To increase the capacity of the replica and for easier removal of the glue the PCB is covered with graphite spray before the wood glue is applied. After the thin layer of half a millimeter wood glue is set, the replica can be carefully peeled of the PCB. The whole process is depicted in Figure 5(a) to 5(d)

4.3 Evaluation

The created forgeries have been successfully tested on all recent flagship mobile phone finger print sensors, which were trained with corresponding original fingerprints. In addition to that we also successfully evaluated them on a legacy sensor from 2004. The forgeries suffice to fool the sensors and it can also be assumed reasonable that the recorded images can be used to track users. This could be done with a small piece of malware that performs the steps taken to extract the fingerprints automatically. By applying one of the well known fingerprint recognition algorithms, for example [13], on that extract a user can be reliably identified and tracked even across devices. With the performance available on modern devices this can be even done on the system itself.



(a) Picture from camera

(b) Extracted negative

Figure 6: By using binary-imaging techniques the image extract as in Figure 4 can be transformed to a negative to be used during fingerprint forgery.

5 Discussion

With the method we presented an attacker can use reflections in the user's face to perform keylogging with a smartphone's front camera. We could successfully demonstrate that the accuracy for this technique outperforms previous methods. An additional feature of our technique is that, in contrast to the work of Xu et al. [27], it can be utilized remotely.

By evaluating the mitigation strategies proposed in that work we could furthermore demonstrate that and why they are not effective in the context of our method. Only a fully randomized keyboard provides some security. However, this approach will either fall to statistical analysis, or the constantly increasing resolution of user facing cameras. The experiments we conducted to verify our technique furthermore demonstrate, that it is not necessary to implement a complex analysis framework in a real-world case, as the input extraction can be as easily conducted manually.

Furthermore we have created a method that allows an attacker to extract its victims fingerprints with a normal smartphones camera in a quality high enough to create usable forgeries. An attacker can use these forgeries to circumvent stationary access controls for secured areas or use them to plant false evidence at a crime scene. With our research we could also demonstrate that it is possible to use these forgeries to circumvent the most advanced sensors for mobile phones recently introduced to the market [25].

This provides an important opportunity during targeted attacks in contrast to traditional fingerprint extraction methods [5]. Possible attackers do not have to to rely on prints being present on an obtained phone. Hence they may have the opportunity to circumvent local fingerprint authentication on a phone to steal confidential data before the victim can issue a remote wipe. Additionally, if the fingerprint itself is used in a biometric remote authentication scheme as first proposed by Boyen et al. [3], such a system is effectively broken by the introduced method, as the cryptographic secrets in that case are bound to the extracted fingerprint.

Finally, a wide range of privacy violations follows from these attacks. Authors of malicious software may use the fingerprints of a user to track the user across multiple devices or distinguish multiple users of one device. These techniques are also relevant, if an attacker has to reliably establish the identity of a user to make sure that the right device has been compromised, for example in case of high profile target.

5.1 Mitigation

The presented attack vectors create severe challenges for a users' security and privacy. As already discussed in the introduction of this paper, permission systems do not provide sufficient protection, as long as a user can grant those permissions to applications asking for them. Therefore we will focus on mitigation strategies that do not require additional decisions from the user.

The most convenient technique imaginable is a dedicated hardware lid, which physically disables the camera. This can either happen with a shutter or by separating the power connection. To enhance the security of this technique, a dedicated sensor in the trusted computing environment indicating the state of the disable button could be implemented, that can be checked by applications. Hence an application could refuse logins, if the camera is not effectively disabled. In fact, such a sensor, disabling all non essential sensory inputs on a device would also effectively mitigate any other sensor side-channel based keylogger method.

The extraction of fingerprints can not be mitigated as easily. While such a button might help in restricting the amount of situations in which a fingerprint can be extracted, it does not prevent all of these situations. Especially if a user forgets to close the lid before putting the phone down. This leaves two possibilities. For high security phones removing the cameras all together is certainly an option, and due to the uncovered issues advisable. For normal end-user, however, it is not. To mitigate the presented attacks on those systems an in-camera algorithm that reduces the resolution of parts of an image that have been identified as fingerprints, before the image leaves the sensor may be applicable. Similar to the lid the power-supply for the camera can be coupled to the one of the screen. Hence if the screen is of, as the phone has been put down, the camera is necessarily off. Other methods include using biometric features that are invisible in normal light, for example deep vein patterns in the human finger.

6 Conclusion

Within this paper we have identified that and how an attacker can exploit the camera of a user's smartphone to obtain sensitive information. We established that the cameras found in modern smartphones constitute a serious security threat. Not only could the camera be abused for a keylogger, it could also be the extraction point for the user's fingerprints. Furthermore we have demonstrated how determined attackers could use these weaknesses to steal sensitive information and penetrate high security environments. We also investigated multiple mitigation strategies, which would prevent attackers from exploiting the camera of a device for keylogging and severely hampered attempts to extract fingerprint information if they were widely adopted. Fully mitigating the extraction of visible biometric features in the presence of a camera is however hardly possible.

This leads to the conclusion that phones used in high security environments should avoid having cameras. As it is considerably hard to effectively re-issue biometric identifiers like fingerprints, this basically holds for every end-user device handled by personnel working in a high security environment. As this is unlikely, the mitigation methods discussed in Section 5 should be implemented.

6.1 Further Work

Although we have demonstrated that a facial reflection based keylogger poses a real threat, an implementation of such a system should be empirically evaluated on various camera resolutions and facial reflection surfaces. As modern phone cameras work only in the visible spectrum of light, the development of sensors and identification of biometric features that can not be recorded in the visible spectrum should be a high priority. Although such methods already exit, the associated sensors are usually to big to be integrated in mobile devices. Changing this should be considered during further research as well.

Acknowledgments

This work was partially supported by the EU FP7/2007-2013 (FP7-ICT-2011.1.4 Trustworthy ICT), under grant agreement no. 317888 (project NEMESYS). We would like to thank the *ct'* Security editorial office for providing us with the OPPO N1 and the anonymous peer reviewers for their valuable feedback.

References

[1] Aviv, A. J., Sapp, B., Blaze, M., and Smith, J. M. Practicality of accelerometer side channels on smartphones. In *Proceedings of the 28th Annual Computer Security Applications Conference* (2012), ACM, pp. 41–50.

[2] Backes, M., Durmuth, M., and Unruh, D. Compromising reflections - or - how to read lcd monitors around the corner. In *Proceedings of the IEEE Symposium on Security and Privacy* (2008).

[3] Boyen, X., Dodis, Y., Katz, J., Ostrovsky, R., and Smith, A. Secure remote authentication using biometric data. In *Advances in Cryptology–EUROCRYPT 2005*. Springer, 2005, pp. 147–163.

[4] Brocker, M., and Checkoway, S. iseeyou: Disabling the macbook webcam indicator led.

[5] CCC. Chaos Computer Club breaks Apple TouchID `http://www.ccc.de/en/updates/2013/ccc-breaks-apple-touchid`, accessed: 06.05.2014, 2013.

[6] Damopoulos, D., Kambourakis, G., and Gritzalis, S. From keyloggers to touchloggers: Take the rough with the smooth. *Computers & Security* (2012).

[7] Felt, A. P., Chin, E., Hanna, S., Song, D., and Wagner, D. Android permissions demystified. In *Proceedings of the 18th ACM conference on Computer and communications security* (2011), ACM, pp. 627–638.

[8] Felt, A. P., Ha, E., Egelman, S., Haney, A., Chin, E., and Wagner, D. Android permissions: User attention, comprehension, and behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security* (2012), ACM, p. 3.

[9] Froomkin, A. M. The death of privacy? *Stanford Law Review* (2000), 1461–1543.

[10] Fujieda, I., and Haga, H. Fingerprint input based on scattered-light detection. *Applied optics 36*, 35 (1997), 9152–9156.

[11] Geller, B., Almog, J., Margot, P., and Springer, E. A chronological review of fingerprint forgery. *Journal of forensic sciences 44* (1999), 963–968.

[12] Jain, A. K., Chen, Y., and Demirkus, M. Pores and ridges: high-resolution fingerprint matching using level 3 features. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 29*, 1 (2007), 15–27.

[13] Jiang, X., and Yau, W.-Y. Fingerprint minutiae matching based on the local and global structures. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on* (2000), vol. 2, IEEE, pp. 1038–1041.

[14] Kamp, K. A., Timmerman, N., Lind, G., Graybill, J., and Natowsky, I. Discovering childhood: Using fingerprints to find children in the archaeological record. *American Antiquity* (1999), 309–315.

[15] Kücken, M. Models for fingerprint pattern formation. *Forensic science international 171*, 2 (2007), 85–96.

[16] KÜCKEN, M., AND NEWELL, A. C. Fingerprint formation. *Journal of theoretical biology 235*, 1 (2005), 71–83.

[17] KUHN, M., AND KUHN, C. Compromising emanations: eavesdropping risks of computer displays. In *Technical report* (2003), University of Cambridge.

[18] LAVIE, A., AND DENKOWSKI, M. J. The meteor metric for automatic evaluation of machine translation. *Machine Translation 23*, 2-3 (2009), 105–115.

[19] LI, J., ZHENG, R., AND CHEN, H. From fingerprint to writeprint. *Communications of the ACM 49*, 4 (2006), 76–82.

[20] MATSUMOTO, T., MATSUMOTO, H., YAMADA, K., AND HOSHINO, S. Impact of artificial gummy fingers on fingerprint systems. In *Electronic Imaging 2002* (2002), International Society for Optics and Photonics, pp. 275–289.

[21] OPPO. OPPO N1 `http://en.oppo.com/products/n1/`, accessed: 20.05.2014, 2014.

[22] RATHA, N. K., BOLLE, R., ET AL. *Automatic fingerprint recognition systems*, vol. 1. Springer, 2004.

[23] SAGIROGLU, S., AND CANBEK, G. Keyloggers. *Technology and Society Magazine, IEEE 28*, 3 (2009), 10–17.

[24] SIMON, L., AND ANDERSON, R. Pin skimmer: Inferring pins through the camera and microphone. In *Proceedings of the 3rd ACM workshop on Security and privacy in smartphones and mobile devices* (2013), ACM.

[25] SIMONITE, T. Pay with Your Fingerprint `http://m.technologyreview.com/news/525996/pay-with-your-fingerprint/`, accessed: 30.04.2014, 2014.

[26] TSIKOS, C. Capacitive fingerprint sensor, Oct. 5 1982. US Patent 4,353,056.

[27] XU, Y., HEINLY, J., WHITE, A. M., MONROSE, F., AND FRAHM, J.-M. Seeing double: Reconstructing obscured typed input from repeated compromising reflections. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security* (2013), CCS '13, pp. 1063–1074.

[28] XU, Z., BAI, K., AND ZHU, S. Taplogger: Inferring user inputs on smartphone touchscreens using on-board motion sensors. In *Proceedings of the fifth ACM conference on Security and Privacy in Wireless and Mobile Networks* (2012), ACM, pp. 113–124.

[29] ZHOU, Y., AND JIANG, X. Dissecting android malware: Characterization and evolution. In *Security and Privacy (SP), 2012 IEEE Symposium on* (2012), IEEE, pp. 95–109.