Biometric Authentication Under Threat: Liveness Detection Hacking

Yu Chen
Tencent Security Xuanwu Lab
alohachen@tencent.com

Bin Ma
Tencent Security Xuanwu Lab
jacksonma@tencent.com

Zhuo Ma
Tencent Security Xuanwu Lab
hyperchemma@tencent.com

July 29, 2019

Abstract

Biometric authentication has been widely used in scenarios such as device unlocking, App login, real-name authentication and even mobile payment. It provides people with a more convenient authentication experience compared with traditional technique like password. A classic biometric authentication process includes biometrics collection, preprocessing, liveness detection and feature matching. With the leakage of biometric data and the enhancement of AI fraud ability, liveness detection has become the Achilles' heel of biometric authentication security as it is to verify if the biometric being captured is an actual measurement from the authorized live person who is present at the time of capture. Previous research mainly focused on how to generate fake data but lack of survey on the security of liveness detection. The purpose of our work is to expose the vulnerability of liveness detection and warn us that once it is compromised, hackers can bypass biometric authentication system easily. In this paper, we will introduce our arsenal of attacking liveness detection and show how to apply them to bypass several off-the-shelf biometric authentication products, including 2D/3D facial authentication and voiceprint authentication.

***K*eywords** Liveness Detection · Facial Recognition · Voiceprint Recognition · FaceID

1 Introduction

With the leap-forward development of artificial intelligence technology, biometric authentication techniques, which rely on the inherited biometric traits taken from the user himself for authentication, have gained wide range of applications recently. Thanks to the convenience of biometric authentication, more and more people tend to use it for mobile phone unlocking, App login, and mobile payment.

A classic biometric authentication process includes biometrics collection, preprocessing, liveness detection and feature matching. With the leakage of biometric data and the enhancement of AI fraud ability, liveness detection has become the Achilles' heel of biometric authentication security as it is to verify if the biometric being captured is an actual measurement from the authorized live person who is present at the time of capture. In recent years, in order to counter the escalating attacks, biometric authentication providers has developed a variety of liveness detection mechanism. For example, texture detection, optical flow detection and attention detection are employed for facial recognition; synthesized speech detection and playback reverberation detection are employed for voiceprint recognition. But can these liveness detection mechanisms be effective against hackers? Unfortunately, previous research mainly focused on how to generate fake data but lack of survey on the security of liveness detection.

The purpose of our work is to expose the vulnerability of liveness detection and warn us that once it is compromised, hackers can bypass the biometric authentication system easily. In this talk, we'll fist introduce our arsenal of attacking liveness detection which includes the following two kinds of weapons:

- Injecting fake video or audio streams by evil hardware to hidden attack media
- Creating specific recognition scene to trigger the defect of liveness detection algorithm

Contributions. The contributions of the paper are out-lined as follows:

- To the best of our knowledge, we are the first to propose a universal methodology for attacking liveness detection in multiple types of biometric authentication.
- We reversed the attention detection mechanism of FaceID and bypass it with ULTRA-LOW COST.

2 Preliminaries

2.1 Biometric Authentication

Biometric authentication is a security process that relies on the unique biological features of an individual to verify that he is who is says he is. Compared with traditional ways like password, pin or token identification, biometric features (including facial, voiceprint, fingerprint, iris, palm veins, palm print, DNA, hand geometry) provide a more convenient and accurate authentication mechanism to identify.

Biometric authentication system compares the captured biometric data with the pre-stored data and finally confirms the results. Howerer, physiological biometric data of individual is hard to modify, which mean if one's biometric data was leaked, he or she can not modify the password.

2.2 Liveness Detection

Biometric authentication is susceptible to "presentation attacks" such as photo spoofing and recording playback. Liveness detection is to verify if the biometric being captured is an actual measurement from the authorized live person who is present at the time of capture. With the leakage of biometric data and the enhancement of AI fraud ability, liveness detection has become the Achilles' heel of biometric authentication security.

2.3 Existing liveness detection methods

2.3.1 Texture-based Methods

Since most face recognition systems adopt only RGB cameras, using texture informationhas been a natural approach to tackling face anti-spoofing. For example, classic LBP, HoG, SIFT, and SURF textures, as well as the recent CNN-based anti-spoofing method using texture.

2.4 Temporal-based Methods

One of the earliest solutions for face anti-spoofing is based on temporal cues such as eye-blinking, the motion of mouth and lip to detect the face liveness. While these methods are effective to typical paper attacks, they become vulnerable when attackers present a replay attack.

2.4.1 Optical Flow

Optical flow is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer and a scene. The difference among optical flow fields can be calculated to determine whether a face is a real face or not.

2.4.2 Attention Detection

Attention detection is designed to confirm user attention and intent to unlock by detecting that your eyes are open and directed at your device. The facial authentication system on most mobile phones has this feature enabled by default.

2.4.3 Playback Reverberation

Reverberation is the collection of reflected sounds from the surfaces in an enclosure like an auditorium. The data of recording voice will lose some information of frequency domain, and is different with the voice of a real person speaking, which can be used to defeat a playback attack.

There are still many ways to perform in liveness detection, but most of them are based on the difference between the attack medium characteristics and the living characteristics.

3 Related Work

Recently, the topic of liveness detection for biometric authentication systems has gained a great deal of interest among biometric researchers in both academia and industry. One typical method is find the difference and features between the living and nonliving things such as color, texture, skin, movement etc. D. Wen[1] proposed a face spoof detection algorithm based on specular reflection, blurriness, chromatic moment, and color diversity extracted from face image. Z. B[2] exploit the joint colour-texture information from the luminance and the chrominance channels by extracting complementary low-level feature descriptions from different colour spaces to detect face spoofing. S. B[3] presents a new framework for face spoofing detection in videos using motion magnification and multi feature evidence aggregation in a windowed fashion. X. L[4] proposed a robust anti-spoofing method by detecting pulse from face videos based on the fact that a pulse signal exists in a real living face but not in any mask or print material.

Besides, many deep learning algorithm have been integrated in liveness detection. Z Xu[5] proposed a deep neural network architecture combining Long Short-Term Memory (LSTM) units with Convolutional Neural Networks (CNN) for face anti-spoofing. G. B. de S.[6] proposed a novel CNN architecture trained in two steps by which each part of the neural network learns features from a given facial region and the whole model is fine-tuned on the whole facial images. Y Atoum[7] extracted the local features and holistic depth maps from the face images for face anti-spoofing. Y Liu[8] proposed a CNN-RNN model to estimate the face depth with pixel-wise supervision, and to estimate rPPG signals with sequence-wise supervision.

There are also some attacks to break biometric authentication system. As early as in 2009, D. N[9] had broke the face recognition authentication system in laptops. J. G[10] proposed a probabilistic approach to reconstruct iris images from binary templates to break the iris recognition systems, but both of them was outdated and new features were adopted to mitigate that, our attack based on hardware injecting can bypass most of the new liveness detection method and finish a remote attack with low effort. Lei Li[11] proposed a novel detection method for 3D face mask presentation attack by modeling reflectance differences based on intrinsic image analysis. The researcher of Bkav[12] from Vietnam crafted a 3D mask, which beats FaceID in the way that twins unlock iPhone X, those attack based on 3D mask sounds available but it requires 3D information of the victim and is proven to be difficult to replicate.

As for voiceprint, University of Alabama at Birmingham researchers[13] have found that automated and human verification for voice-based user authentication systems are vulnerable to voice impersonation attacks. Zhang[14] design a totally inaudible attack called DolphinAttack which modulates voice commands on ultrasonic carriers to achieve inaudibility by leveraging the nonlinearity of the microphone circuits. Yuan[15] found that the voice commands can be stealthily embedded into songs which can effectively control the target system through ASR without being noticed. John Seymour and Azeem Aqil[16] use freely available machine learning models to spoof voice authentication algorithms in Blackhat USA 2018. However, John's research mainly focuses on the generation of arbitrary voice. But in the actual attack scenario, we need not only the synthesis of voice but also bypassing the liveness detection algorithm. In this paper, we bypass the financial-level liveness detection algorithm by audio injection. Meanwhile, most Apps use digital voiceprint locks for authentication which means that we don't need to generate arbitrary voice but only the voice of the ten digits 0-9. Therefore, we propose a simpler voice synthesis method that requires only a short phone recording as a corpus.

4 Methodology

In this paper, we design, implement and evaluate a series of methods for hacking liveness detection. This section will summarize the two main ideas of our methodology.

4.1 Injecting fake biometric streams by evil hardware to hidden attack media

Injecting fake biometric streams such as video and audio by evil hardware have the following three benefits:

- Avoid information loss during biometric secondary acquisition and playback, such as HSL space color loss, focus blur, playback reverberation effect, etc.
- Since the attack medium is not exposed, the attack mode can bypass the liveness detection algorithm which based on the medium characteristics, such as texture, optical flow, frequency response distortion, etc.
- Since current biometric authentication system cannot verify the validity of the sensor-level data stream, the attack mode is completely software-insensitive.

4.2 Creating specific recognition scene to trigger the defect of liveness detection algorithm

In order to improve the recall rate of authentication under certain special recognition scenes (such as face recognition in a sunglasses scene and voiceprint recognition in a noisy environment), current liveness detection algorithm must be at the cost of increasing the FPR (False Positive Rate). Through a lot of experimental observations, we found that attacker can bypass liveness detection by deliberately creating these malicious scenarios that can trigger algorithmic defects.

5 Video Injection Example

The Camera Serial Interface (CSI) is a specification of the Mobile Industry Processor Interface (MIPI) Alliance. It defines an interface between a camera and a smartphone processor. We can use Toshiba's TC358749XBG chip to make a hardware module that can converts HDMI stream to MIPI CSI stream. Then we connect the module to an Android development board (RK3399) to form a complete video injection attack device (Figure 1). Using this device, we can disguise the hdmi output of a PC or notebook as a video stream captured by an Android device camera. In this way, we can not only avoid information loss in the secondary acquisition of the face such as HSL space color loss and focus blur, but also make the playback process transparent, that is, without introducing texture and optical flow characteristic of paper or screen.

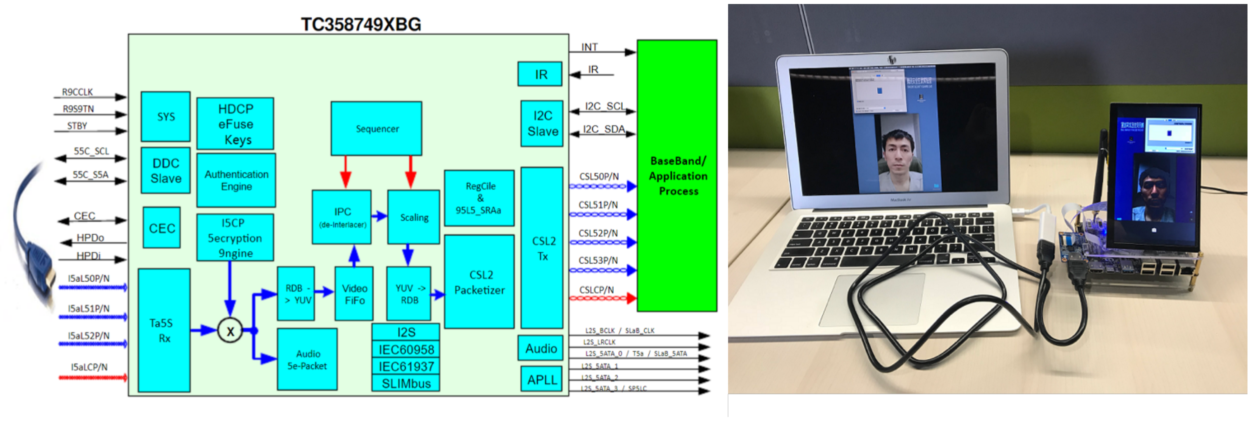


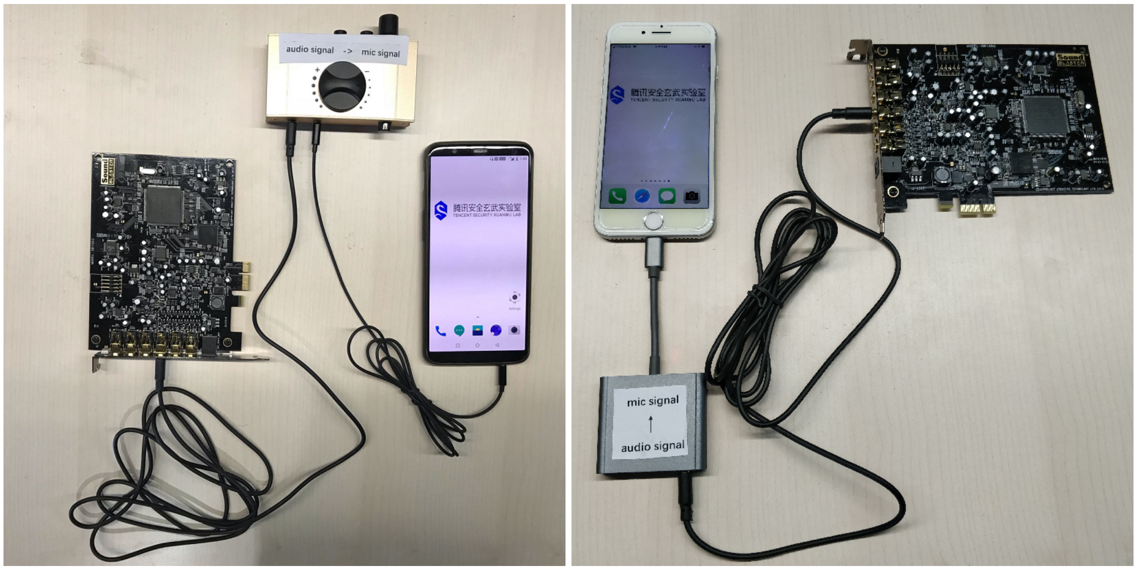
Figure 1: Video injection device based on TC358749XBG

6 Audio Injection Example

We found that most voiceprint authentication systems can receive voice input from the microphone cable. A microphone-level signal is the voltage level that comes out of a microphone when someone speaks into it, typically just a few ten-thousandths of a volt. This voltage varies in response to changes in voice level and and in the talker-to-mic distance. But the signal is still quite small. We can use two types of analog circuits (Fig. 2) to create a hardware module that converts the audio stream into a microphone stream, one for the Andriod device and the other for the iOS device. Then we connect it to an real phone or tablet to form a complete audio injection attack device. With this device, we can directly convert the audio stream of the headphone output port into a audio stream of the microphone input port and directly inject malicious voice into the device. In this way, we can not only avoid the loss of information during audio playback, but also make the playback reverberation effect and frequency response distortion disappear.

7 Bypass FaceID's attention detection

With a simple glance, Face ID securely unlocks iPhone X. It provides intuitive and secure authentication enabled by the TrueDepth camera system, which uses advanced technologies to accurately map the geometry of your face. Face ID confirms attention by detecting the direction of your gaze, then uses neural networks for matching and anti-spoofing so you can unlock your phone with a glance [17]. In this chapter, we will reverse the attention detection mechanism of FaceID and show how to use glasses and tape to unlock the FaceID while victim is sleeping.



(a) For Android Devices

(b) For iOS Devices

Figure 2: Two types of audio injection device

7.1 Eyes abstraction in dark light

To enhance the user experience, FaceID allows users to unlock while wearing sunglasses. However with sunglasses, the light at the eyes is diminished. We use the infrared camera experiment to find the abstraction of the eyes in low light environment as shown in Figure 3. In the dark environment, the abstraction of the eye is a black area with a white point in the center. We also found that after identifying the glasses, FacID does not extract 3D information from the eye area.

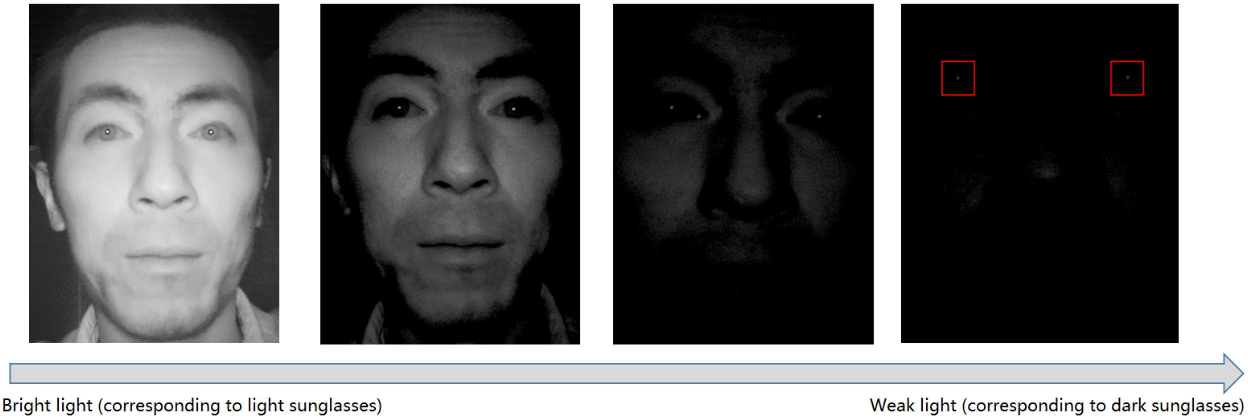

Bright light (corresponding to light sunglasses)
Weak light (corresponding to dark sunglasses)

Figure 3: Eyes abstraction in dark light

7.2 The relationship between white spot position and gaze direction

The relationship between white spot position and gaze direction is as shown in Figure 4. For example, if eyes looking upwards, the white spot is in the lower part of the black area. Conversely, if eyes looking forward, the white spot is in the center part of the black area. For example, if the eye is looking up, the white spot is in the lower part of the black area. When the user normally unlocks the phone, the eye looks forward and the white spot is in the center of the black area.

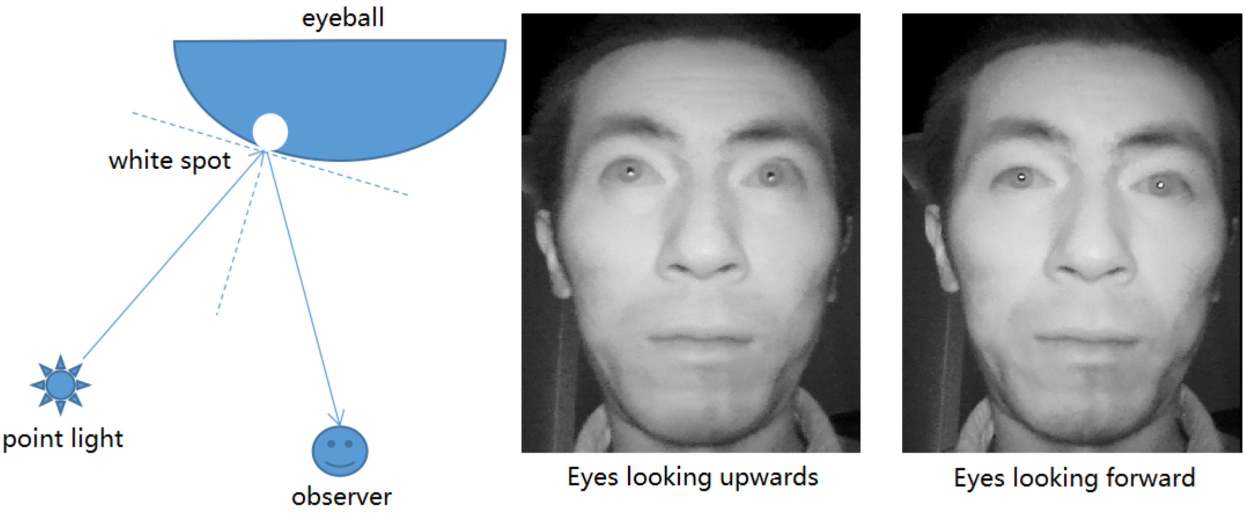


Figure 4: The relationship between white spot position and gaze direction

7.3 The Prototype of X-glasses

Through the above analysis of the attention detection mechanism, we propose the X-glasses prototype shown in Figure 5. The only material needed to make X-glasses is black and white tape and a pair of ordinary glasses. The entire production process takes no more than two minutes and does not require customization for specific victims. The black tape simulates the pupil of the eye, and the white spot simulates the reflection of the infrared supplement light source from the mobile phone. The idea behind the X-glasses is to use eyeglass frames to switch the eye recognition mode to simple mode, as we summarized in the Section 4.2.

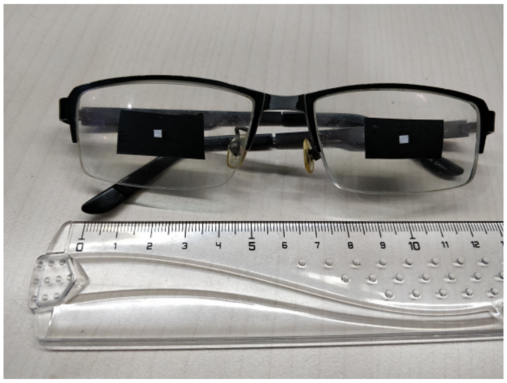


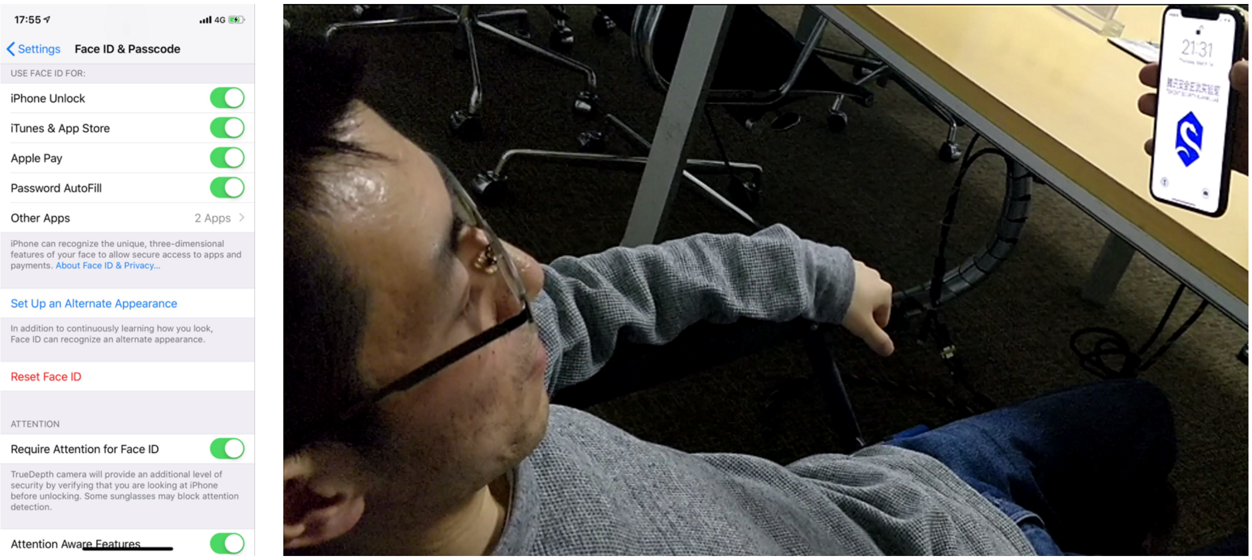
Figure 5: The Prototype of X-glasses

7.4 Bypass FaceID's attention detection

The attack scenario is shown in Figure 6. In fact, the glasses can be held by the hand so that it does not touch the victim. In this demo we unlock victim's mobile phone and then transfer his money through mobile payment App.

8 Mitigation Measures

Most of authentication system collect biometric data by various hardware or sensor, but failed to verify whether the hardware is trusted or not. The secure biometric authentication system should check every process of the chain, including the data input or collecting node to avoid hardware injection. For example, checking the brand, identifier or other hardware info to verify its authenticity, or construct a hardware fingerprint to against hardware spoofing.



17:55 4G

Settings Face ID & Passcode

USE FACE ID FOR:

iPhone Unlock ☑

iTunes & App Store ☑

Apple Pay ☑

Password AutoFill ☑

Other Apps 2 Apps >

iPhone can recognize the unique, three-dimensional features of your face to allow secure access to apps and payments. About Face ID & Privacy...

Set Up an Alternate Appearance

In addition to continuously learning how you look, Face ID can recognize an alternate appearance.

Reset Face ID

ATTENTION

Require Attention for Face ID ☑

TrueDepth camera will provide an additional level of security by verifying that you are looking at iPhone before unlocking. Some sunglasses may block attention detection.

Attention Aware Features ☑

Figure 6: Bypass FaceID's attention detection

The server side should strengthen the defense to against further attack. Since the fake merged biometric data need to be generated from the origin biometric data of victim, there are some features to identify and distinguish. For instance, merged voiceprint stream may have some pieces of uniformly distributed voice or irregular curve between two sound waves. Video synthesis stream is usually generated from a photo of victim, which may contain some unsmooth switch at the part of eye or mouth, or different background in all frames on the whole capture process. Besides, there are some deep learning algorithms can detect the synthesis media stream, which is a good choice to conduct a protection mechanism on the server side.

We do not recommend enabling biometric-based account login or password retrieval on unauthorized devices because the hardware of such devices is easily replaced and then injected with faked biometrics. Many app use IMEI, Android ID or MAC address to generate a device fingerprint to bind device with biometric authentication system, but we find it can be bypass with little reversing work. Vendors should protect their codes to generate a device fingerprint or design a device binding mechanism to against device ID spoofing attack.

The reason why X-glass attacks can be successful is that FaceID cannot distinguish between real eyes and fake eyes when users wear glasses. We recommend that the texture features and depth information of the eye be used as a basis for judging whether the eye exists, so as to effectively resist the X-glasses attack while ensuring the recall rate.

References

[1] Di Wen, Hu Han, and Anil K Jain. Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security*, 10(4):746–761, 2015.

[2] Zinelabidine Boulkenafet, Jukka Komulainen, and Abdenour Hadid. Face spoofing detection using colour texture analysis. *IEEE Transactions on Information Forensics and Security*, 11(8):1818–1830, 2016.

[3] Talha Ahmad Siddiqui, Samarth Bharadwaj, Tejas I Dhamecha, Akshay Agarwal, Mayank Vatsa, Richa Singh, and Nalini Ratha. Face anti-spoofing with multifeature videolet aggregation. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 1035–1040. IEEE, 2016.

[4] Xiaobai Li, Jukka Komulainen, Guoying Zhao, Pong-Chi Yuen, and Matti Pietikäinen. Generalized face anti-spoofing by detecting pulse from face videos. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 4244–4249. IEEE, 2016.

[5] Zhenqi Xu, Shan Li, and Weihong Deng. Learning temporal features using lstm-cnn architecture for face anti-spoofing. In *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pages 141–145. IEEE, 2015.

[6] Gustavo Botelho de Souza, João Paulo Papa, and Aparecido Nilceu Marana. On the learning of deep local features for robust face spoofing detection. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 258–265. IEEE, 2018.

[7] Yousef Atoum, Yaojie Liu, Amin Jourabloo, and Xiaoming Liu. Face anti-spoofing using patch and depth-based cnns. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 319–328. IEEE, 2017.

[8] Yaojie Liu, Amin Jourabloo, and Xiaoming Liu. Learning deep models for face anti-spoofing: Binary or auxiliary supervision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 389–398, 2018.

[9] Nguyen Minh Duc and Bui Quang Minh. Your face is not your password. In *Blackhat USA*, 2009.

[10] Javier Galbally, Arun Ross, Marta Gomez-Barrero, Julian Fierrez, and Javier Ortega-Garcia. From the iriscode to the iris:a new vulnerability of iris recognition systems. In *Blackhat USA*, 2012.

[11] Lei Li, Zhaoqiang Xia, Xiaoyue Jiang, Yupeng Ma, Fabio Roli, and Xiaoyi Feng. 3d face mask presentation attack detection based on intrinsic image analysis. *arXiv preprint arXiv:1903.11303*, 2019.

[12] Bkav. Bkav's new mask beats face id in "twin way": Severity level raised, do not use face id in business transactions. `http://www.bkav.com/`. Nov 27, 2017.

[13] Dibya Mukhopadhyay, Maliheh Shirvanian, and Nitesh Saxena. All your voices are belong to us: Stealing voices to fool humans and machines. In *European Symposium on Research in Computer Security*, pages 599–621. Springer, 2015.

[14] Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. Dolphinattack: Inaudible voice commands. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 103–117. ACM, 2017.

[15] Xuejing Yuan, Yuxuan Chen, Yue Zhao, Yunhui Long, Xiaokang Liu, Kai Chen, Shengzhi Zhang, Heqing Huang, Xiaofeng Wang, and Carl A Gunter. Commandersong: A systematic approach for practical adversarial voice recognition. In *27th {USENIX} Security Symposium ({USENIX} Security 18)*, pages 49–64, 2018.

[16] John Seymour and Azeem Aqil. Your voice is my passport. In *Blackhat USA*, 2018.

[17] Apple. Faceid security guide. `https://www.apple.com/business/site/docs/FaceID_Security_Guide.pdf`. Accessed November, 2017.