Biometric PAD Toolbox Overview [Toolbox]

Version 1.2, October 15, 2024

Table of Contents

Chapter 1. Introduction

The TOE may be vulnerable to presentation attacks where attackers attempt to subvert the biometric enrolment or verification by presenting the Presentation Attack Instruments (PAIs). There is a wide range of PAIs that can be used, including natural biometric characteristics, such as dead eyes, or artefacts created from copied or faked characteristics. Using natural biometric characteristics is out of scope of [BIOPP-Module] evaluation and the evaluator shall only use created artefacts to evaluate the TOE.

The toolbox defines the common artefacts for each biometric modality based on publicly available information (e.g. research papers), experiences and knowledge shared among the BIO-iTC members. The evaluator needs to read the [BIOSD] Section 7 as it explains how the evaluator shall use the toolbox during the ATE_IND.1 (Independent testing) and AVA_VAN.1 (Penetration testing) evaluation for Presentation Attack Detection (PAD) in detail. In this evaluation, PAD is being measured using Imposter Attack Presentation Attack Rate (IAPAR), which is a full system review as opposed to a component level review which would evaluate the matching subsystem and PAD subsystem separately.

This overview is originally developed for evaluation activities for FIA_MBV_EXT.3, however, the evaluator can apply the same principles to evaluation activities for FIA_MBE_EXT.3.

Chapter 2. Biometric modalities supported

Currently toolbox is developed for the following biometric modalities:

- Eye
- Face
  - 2D Image
  - 3D Image
- Fingerprint
- Vein

The toolbox is intended to be a state of the art set of test and is continually updated. The latest versions of the toolbox for each supported modality must be used by the evaluator at the time the testing begins.

Chapter 3. Purpose of this toolbox overview

This toolbox overview describes common instructions and components for all toolboxes, rather than repeating the same information in each toolbox. The evaluator shall refer both the toolbox overview and the relevant toolbox to perform independent and penetration testing. If there is conflict between the toolbox overview and a toolbox, the toolbox takes precedence over the overview.

Chapter 4. Structure of the Toolbox

Each toolbox shares the same structure and includes following sections.

***modality name* Toolbox overview**

This section provides specific information only applicable to relevant biometric modality.

***modality name* Toolbox Inventory**

This section categorizes tools and materials that the evaluator shall use to capture an image of biometric characteristics and produce artefacts.

***modality name* Verification List**

This section summarizes all test items that the evaluator shall perform during independent testing. As explained in [BIOSD] Section 7, the evaluator shall select specific test items for penetration testing based on the result of independent testing.

***modality name* References**

This section lists all publicly available information referred to create a toolbox.

Test items

Each test item includes the following sub-sections. This toolbox overview provides a general test protocol in common for all toolboxes and these test items describe more detailed information to enable repeatable testing.

Table 1. Test Items

Sub-section name	Description
Number	Identification number of test
Attack type	Category of attack
Overview	General overview of test
Input	Required input to produce artefacts
Attack Tools/Media	Required tools and media to capture an image of biometric characteristics and produce artefacts
Recipe	Procedure to create artefacts
Variations	Variants of artefacts to be generated based on this test item. The evaluator shall create those variants by slightly different procedure (e.g. different **Recipe** or with different **Attack Tools/Media** specified here) for the independent testing.
Prerequisite	Any conditions that should meet to perform each test
Presentation	Instructions to present artefact to the TOE

Sub-section name	Description
Penetration Testing and Attack Potential Rating Suggestions	Suggestions that the evaluator should consider devising penetration tests from this test item and calculate the attack potential rating. The evaluator may change the rating considering actual expertise or knowledge of TOE used to succeed attacks, however, the evaluator shall report such changes with proper justification
Pass Criteria	Additional information for pass-fail criteria of the test (IAPAR shall not exceed the value assigned in FIA_MBV_EXT.3.1 in any case)

Chapter 5. Finding and accessing the Toolboxes

To provide flexibility in support for testing the various biometric modalities for PAD, the versioning of the Toolboxes are maintained independently from the versioning of the primary documents ([BIOPP-Module] and [BIOSD]). Each Toolbox is maintained separately within its own GitHub repository so updates targeted to specific modalities can be updated independently as needed over time.

There are several ways to find the Toolboxes that are available.

The simplest method is to go to the GitHub Public Release Packages table from the Biometrics Security iTC homepage. Each available Toolbox will be listed along with a direct link in GitHub to the most recent toolbox package for that modality.

The second method is to go to the Biometrics Security iTC organization in GitHub and find the repositories for each modality there. Each modality has a repository titled <Modality>-Toolbox (where <Modality> would be replaced by a supported modality type). From the home page of the repository, on the right side there is a section titled "Releases". Here you will find all the released versions of the particular toolbox.

5.1. Toolbox versioning

To keep the versioning simple, each released update is just given a sequential whole number, so 1, 2, 3… (the original release was versioned 1.0, but subsequent updates are following the whole number sequencing).

5.2. Toolboxes and technical decisions

Unlike the primary documents (such as the [BIOPP-Module] and [BIOSD]), toolboxes are always fully updated to the next revision; there are no Technical Decisions applied to a Toolbox, it is updated, approved and released as a new version (i.e. moved from v2 to v3).

5.3. Choosing the correct version

In general it is expected that an evaluation will utilize the most recent version of the Toolbox as of the time the evaluation was started (as defined by the scheme). As the Toolboxes can be updated at any time, the evaluation start date is used to help vendors freeze the requirements for their products.

It is possible for a scheme to have different requirements about what version of a Toolbox should be used, which supercedes any recommendations made by the iTC.

5.4. Referencing toolboxes in an evaluation

As all evaluations must properly reference the Protection Profiles and Supporting Documents, the

Toolbox(es) used in an evaluation claiming support for PAD must list the specific versions of any Toolboxes.

The reference in the Conformance Claims section of the Security Target should provide the following information to unambigiously point to the correct Toolbox as part of the claims for the PP-Configuration (using the Face Toolbox as an example):

Toolbox: Face Toolbox, Version 2, November 11, 2021 (https://github.com/biometricITC/Face-Toolbox/releases/tag/v2)

All components, including the GitHub link to the specific version must be included in the Security Target.

Chapter 6. Presentation Attack Levels

Biometric presentation attacks are not all rated the same in terms of their attack potential. This reflects the reality that not all biometric systems can resist all types of attacks. To support this while allowing systems that can protect against stronger attacks to highlight additional capabilities, attacks that target higher attack potentials than AVA_VAN.1 are included as part of the toolboxes.

PAD levels are defined based on the tables in [BIOSD] Section 11, and attacks will be grouped based on the modality and PAD level. PAD Levels may be added to any modality

Table 2. PAD Level Attack Potential

PAD Level	**Protected to Values**	**Attack Potential Required to Exploit**
Level 1	< 10	Enhanced Basic
Level 2	10-19	Moderate
Level 3	20-29	High

Higher PAD Levels are inclusive of lower PAD Levels, so a set of tests at PAD Level 2 would incorporate all PAD Level 1 tests. This allows for a reduction in the duplication of attack documentation as test procedures change over time.

As the [BIOPP-Module] relies on the AVA_VAN level specified in the Base-PP, it is possible that the AVA_VAN level for the PAD test may not match the Base-PP. In this case it is up to the evaluator and scheme to determine how to handle the different level claims. The BIO-iTC has determined that there is value in providing PAD tests at higher AVA_VAN levels regardless of the claims by the device that may incorporate the biometric system.

However, the BIO-iTC will explore the possibility of developing a PAD evaluation method at any PAD level that can be mutually recognized under the CCRA, for example, using a multi-assurance evaluation concept that enables the evaluation to conform to a multi-assurance security target.

Chapter 7. Common guidance for Independent & Vulnerability Testing

As explained in [BIOPP-Module], the TOE is the whole biometric system, including Comparison, Decision and Presentation Attack Detection Subsystems. This means in order to successfully overcome the TOE by the use of artefacts, a genuine person (test subject) has to be enroled into the TOE, artefacts have to be created referring the toolbox for the corresponding biometric modality and artefacts have to produce an attack presentation match (i.e. a successful presentation attack).

For all types of testing, there are some common steps/procedures to be followed. These are detailed here.

7.1. General test protocol

Presentation attacks can be performed through the following three steps.

7.1.1. Preparation

Before testing can start, the following pre-requisite needs to be met:

- It has to be ensured that the test subject whose body part is used to produce the artefacts for testing is enroled into the TOE correctly as follows.
  - Enrolment shall be done following guidance provided by the TOE.
  - At least 5 test enrolment transactions shall be performed by the test subject to ensure that the test subject can enrol correctly and be verified after enrolment.
  - In case of repeated failures during the test enrolment, the test subject shall use a different body part (this could mean to use a different finger of the test subject in case of fingerprint verification) and start test enrolment transactions again.
  - If the test subject cannot enrol any body parts during the test enrolment, the test subject shall be exempt from further testing.

7.1.2. Artefact production

Artefact production needs to follow these requirements:

- The evaluator shall document any necessary information so that artefacts used for the test can be re-produced by the evaluator.
- Each produced artefact shall be identified by a unique identifier. This identifier shall be attached to the artefact at all times (as far as this is possible without destroying the artefact).

As the testing described in the individual toolboxes encompasses the creation and presentation of a large number of artefacts, the test report shall provide sufficient information to ensure how the artefacts were created, presented and as applicable, stored (or retrieved from storage).

If the scheme provides a guidance on the level of detail of the report, the evaluator must follow

such guidance. However, if there is no guidance available from the scheme, the BIO-iTC recommends the inclusion of visual (pictures and/or video) evidence in the test report. If sound is included as part of the biometric system, then audio evidence should also be included with the visual evidence.

Broadly, visual/audio evidence should be used on a per-artefact type basis, such that each type is shown clearly once, and the remainder of artefact production and usage would be recorded as expected (but not captured with visual or audio recordings). The evidence collected does not need to be continuous (for example a full video recording of every step), but must record significant steps in the creation of the artefact.

If visual/audio evidence is being provided, the following categories should have visual/audio evidence:

- Creation of an artefact type (significant steps)
- Use of an artefact type (preparation that may be needed, usage)
- If applicable, storage of an artefact type after use
    - How the artefact will be stored for later use
- If applicable, retrieval of an artefact type from storage for use
    - How is the artefact prepared for use after removed from storage

From a planning standpoint, the easiest way to handle this would be to record one artefact from retrieval/production to disposal/storage (depending on the type).

Artefact storage

It is widely known that, for any biometric modality, some degree of variation in the biometric features will occur over time. For example, the levels of skin dryness is different between summer and winter, and captured fingerprint images from the same person may also vary a little but such little difference may affect the biometric performance. So, fingerprint artefacts created in winter should not be used for presentation attacks against enrolment images captured in summer. The quality of artefacts may also be changed over time. For example, glue used for fingerprint artefacts begins to dry and harden and the success rate of attacks may begin to drop within a few weeks after the creation of the artefacts.

The evaluator may reuse artefacts for later evaluations, however, the evaluator shall check that the following conditions are met to reuse the stored artefacts:

- Time difference between enrolment and artefact creation and usage should be as minimal as possible, though individual biometric modalities may have different time periods for which reuse is acceptable. If the evaluator uses artefacts older (or for longer) than one month (here defined as a five week period from enrolment and creation), the evalutor shall follow the guidance in Use of stored artefacts for proper storage and retrieval of the artefacts.
- Artefacts should be properly stored according to the manufacturer's recommendations and remain free of visible defects. Some of these may be obvious, like the proper storage of photographs, but others may be more detailed, requiring temperature and humidity controls. For artefacts that are capable of being stored, information about what is done to store the

artefacts (supported by visual evidence and documentation) along with guidance from the manufacturer that supports the methods implemented.

Use of stored artefacts

For artefacts where long term storage of more than one month and re-use is more subjective (such as the fingerprint artefacts), information about how it was determined whether the artefact was in acceptable condition must be provided (for example levels of dryness and hardness).

If artefacts stored more than one month are used in later evaluations, creation date of the artefacts, number of stored artefacts used and the method of storage must be included with the new evaluation to show proper procedures were followed for handling the artefacts (the method of creation of the stored artefacts does not need to be included).

Before use, the evaluator shall check any stored artefacts for visible changes between the artefact and the subject to determine if the artefact is still acceptable for use. For example, a fingerprint artefact where the subject may have cut on their finger at the time of testing would lead to an artefact not being of sufficient quality.

However, as artefacts may degrade over time in ways that are not visible to the human eye but which may impact PAD performance, a PAD test must not rely solely on stored artefacts. To ensure that artefacts are not failing solely due to some sort of unseen degradation, if stored artefacts are to be used, a PAD test sequence must utilize a combination of both stored and freshly created artefacts.

The purpose of this mix is to provide a check that the stored artefacts by comparing the performance of the stored artefacts to the new ones. Stored artefacts that seem to perform significantly lower then they should must be discarded and fresh artefacts created to replace them. The test report must denote which artefacts were stored vs fresh. It is always recommended to create new artefacts for every test to avoid this check if the time and cost of creation of artefacts is low.

As different modalities have different types of artefacts, the modality overview documents specify the requirements for artefact reuse in terms of the maximum percentage of stored artefacts of a type which can be used in testing.

7.1.3. Presentation of artefacts

The results of the presentation of artefacts is defined as:

Table 3. Artefact presentation results

Result	Definition
Successful (Match) Attack	The TOE matches the artefact to the enroled user
Failed Attack	The TOE rejects the artefact

Chapter 8. Guidance for Independent Testing (ATE_IND.1)

For independent testing, this guidance is common for all toolboxes. More specific guidance for a specific biometric modality is provided in each toolbox.

This is in addition to guidance in Chapter 7, *Common guidance for Independent & Vulnerability Testing*.

8.1. General test protocol

The presentation attack can be performed through the following two steps after performing Section 7.1.1, “Preparation”.

8.1.1. Artefact production

The production of artefacts for each toolbox shall be performed as follows:

- The evaluator shall produce artefacts according to the Verification List defined in the toolbox based on the sensor type. If the sensor type does not match one explicitly listed, then all artefacts must be created (as defined by the Other type).
- The evaluator shall follow instructions in the toolbox to produce artefacts, especially the evaluator shall use tools or materials (e.g. camera, display or printer) that meet requirements in toolbox.
- The evaluator shall produce three artefacts from each test subject.

8.1.2. Presentation of artefacts

The evaluator shall present artefacts to the TOE to perform presentation attacks.

- Each artefact shall be presented to the TOE 10 times

8.2. Number of Subjects

The evaluator shall prepare three test subjects for the above test. A test subject is defined as one individual, and not different body parts from one person (i.e. three fingers from one person could not be considered to be three test subjects for the creation of artefacts).

8.3. Pass/Fail Criteria

A TOE passes the test if and only if it reliably defeats the use of **all artefacts (i.e. 3 X 3 = 9 artefacts in total)** that have to be built according to the toolbox. This means that none of the artefacts must be able to reproducibly overcome the TOE.

To reproducibly overcome the TOE by the use of a **certain artefact** in the outlined test scenario is defined as follows:

Table 4. Pass/Fail Criteria

Attempts/Transactions	**Number of matches**	**Outcome**
10	0	TOE passes this artefact
10	1	TOE passes this artefact
10	2	Additional ten (10) attempts/transactions shall be made
20	2	TOE passes this artefact
Up to 20	3 or more	TOE fails this artefact

As explained in [BIOPP-Module], presentation attacks can be done by attempts or transactions.

The maximum number of attempts/transactions allowed with one artefact is twenty (20). If three (3) matches are made to the artefact, the independent test fails (further attempts/transactions are not necessary even if 20 total attempts/transactions have not yet been made) because the IAPAR has exceeded 15%, the allowable maximum value specified in FIA_MBV_EXT.3.1.

Chapter 9. Guidance for Penetration Testing (AVA_VAN)

The evaluator moves to penetration testing only if the TOE passes independent testing. As described in [BIOSD] Section 7, the evaluator shall select those artefacts that show a higher IAPAR during independent testing or higher quality artefacts.

This is in addition to guidance in Chapter 7, *Common guidance for Independent & Vulnerability Testing*.

9.1. Guidance for AVA_VAN.1

9.1.1. General test protocol

Presentation attack can be performed through the following two steps after performing Section 7.1.1, “Preparation”.

Artefact production

The production of artefacts for each toolbox shall be performed as follows:

- The evaluator should select artefacts in a toolbox that may produce attack presentation match at higher probability considering the result of independent testing.
- The evaluator may refine the production process of artefacts, as explained in [BIOSD] Section 7. The toolbox describes generalized process to produce artefacts referring to research papers. These research papers may describe more detailed information to produce better artefacts. Such information is valuable if the TOE’s PAD algorithm is the same or similar to ones tested by researchers. The evaluator shall consider relevant research papers to be authoritative over the generalized descriptions provided in a toolbox for improving the creation of artefacts.
- The evaluator may produce an arbitrary number of artefacts from each test subject within allowed time period. As described in [BIOSD], penetration testing shall be finished within one week.

Presentation of artefacts

The evaluator shall present artefacts to the TOE to perform presentation attacks.

- Each artefact shall be presented to the TOE an arbitrary number of times within allowed time period. As described in [BIOSD], penetration testing shall be finished within one week.

9.1.2. Number of Subjects

If the evaluator can create artefacts that produce an attack presentation match during independent testing, the evaluator should select the test subjects whose artefacts had successful matches and increase the number of attempts/transactions. The evaluator may replace the test subject for penetration testing as described in [BIOSD] Section 7.

9.1.3. Pass/Fail Criteria

As described in [BIOSD], penetration testing shall be finished within one week. The evaluator may select one or two artefacts and perform an arbitrary number of attempts/transactions within this time period. If the evaluator can create artefacts that reproducibly cause the TOE to achieve an IAPAR higher than what is specified in FIA_MBV_EXT.3.1, the TOE fails AVA_VAN.1 evaluation.

9.2. Guidance for AVA_VAN.2 and Higher

As some PAD toolboxes are targeted to levels above AVA_VAN.1, there will be additional information regarding the guidance about penetration testing. The AVA_VAN.1 guidance is used as the basis for all penetration testing, with additional information specific to that toolbox being provided in the PAD toolbox overview.

ISO/IEC 19989 “Information security — Criteria and methodology for security evaluation of biometric systems” provides such additional information. However, this standard is not specific to the mobile devices and not all evaluation activities may be applicable to mobile device biometric systems.

Chapter 10. Related Documents

- [BIOPP-Module] collaborative PP-Module for Biometric enrolment and verification - for unlocking the device -, September 12, 2022, Version 1.1
- [BIOSD] Supporting Document Mandatory Technical Document: Evaluation Activities for collaborative PP-Module for Biometric enrolment and verification - for unlocking the device -, September 12, 2022, Version 1.1

Chapter 11. Revision History

Table 5. Revision history

Version	Date	Description
0.3	May 30, 2019	Public Review Draft 1
0.5	December 20, 2019	Public Review Draft 2
0.6	March 13, 2020	Proposed Release
1.0	May 11, 2020	Public Release
1.1	September 12, 2022	Update based on changes to the PP-Module v1.1
1.2	October 15, 2024	Update to support new PAD test levels