HECTOS

Harmonized Evaluation, Certification and Testing Of Security products

D4.3
Results of test and evaluation for the selected biometric case studies
EXECUTIVE PUBLISHABLE SUMMARY

NPL
Safran Identity & Security
Fraunhofer IGD

Date: 2017-05-10
Project No: 606861
FOI Designation No: FOI-2015-1958
Dissemination Level: PU(Summary)
Total No of Pages: 6

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 606861.

HECTOS

D4.3
Results of test and evaluation for the selected biometric case studies
EXECUTIVE PUBLISHABLE SUMMARY

Version:	1.0
FOI designation no:	FOI-2015-1958
Responsible:	NPL
Author(s):	Tony Mansfield, Aruna Shenoy (NPL) Pierre Gacon (Safran Identity & Security) Olaf Henniger (Fraunhofer IGD)
Number of pages:	6
Dissemination level:	*PU – valid for the Executive summary*
Start date of project:	Sep, 2014
Duration:	41 months

Executive Summary

HECTOS is a European project focusing on harmonization of evaluation, certification and testing of physical security products. Physical security equipment and systems are very diverse in technology, concept of operation, application area and performance, and similar security products are difficult to compare in terms of performance, accuracy, usage, trust and validation of functionality. Currently, there are very few test, evaluation and certification procedures in Europe that are mutually recognized by different Member States. This leads to fragmentation of the market, as identified in the recent EC Communication on Security Industrial Policy, with negative impacts on both suppliers and users.

The HECTOS project focuses on the evaluation and certification schemes for physical security products, and studies how existing schemes used in other areas could be applied, adapted or developed for products used for physical security of people, property and infrastructure. To analyse, develop, enhance, and experimentally validate evaluation and certification schemes, HECTOS conducts case studies in two priority areas: (i) biometrics and (ii) detection of weapons and explosives.

HECTOS D4.3 analyses and assesses the evaluation and certification processes and outcomes for the biometric case study topic against the elements of the prototype certification scheme outlined in HECTOS D3.2, making recommendations to enhance the viability of harmonized certification schemes for biometric security products.

Key aspects of certification schemes being examined include:

- development and agreement of conformity assessment certification criteria across multiple technologies (e.g., both legacy and new biometric technologies)
- harmonization of requirements between related applications
- inter-laboratory comparison, in particular identifying suitable procedures to ensure repeatability for adversarial tests (i.e., presentation attack resistance / presentation attack detection in the case of biometrics)
- consideration of the possibility for certified evaluations, independent of conformity assessment with quantified performance requirements, or thresholds, and
- consideration of the requirements for accreditation (e.g. interlaboratory comparison, publication of test results) to enable re-use and mutual recognition of evaluation results.

The findings from each topic are sumarised below.

Issues and findings from the evaluation and certification of the image quality of contactless fingerprint sensors

The image requirements based on the current image quality requirements and methodology for an acquisition area were adapted to an acquisition volume in order to adjust and apply this methodology to a contactless fingerprint product. After the self-testing phase achieved satisfactory results, a Personal Identification Verification (PIV) certification process was then engaged with MITRE and FBI.

- Allowing self-testing by supplier in a certification scheme can bring several advantages such as accelerated testing phase particularly for products with short product development lifecycles.
- Results from different test houses help validate evaluation methodology (i.e., confirmation that the measures correlate to accurate recognition performance).
- Provisions for extendibility of an existing certification scheme or adaptation of existing requirements and methods to cover new technology or types of products would provide evidence of the new product meeting the standards in the field.
- Self-testing is dependent on the trustful relationship between the supplier and the certification body.
- Results and extended test methods were reviewed and conformance attested.

Issues and findings from the evaluation of the presentation attack detection and presentation attack resistance of a biometric fingerprint verification systems, and the potential for certification

In order to investigate the mechanisms for harmonized methodologies for Presentation Attack Detection (PAD) of fingerprint recognition systems, the main activities has been:

- Independent vulnerability analysis of the Target of Evaluation (TOE),
- Adversarial penetration tests to determine whether the TOE is resistant against presentation attacks
- Assessment of the attack potential and
- Interlaboratory comparison of results, requirements, methods repeatability and consistency of results.

Findings and issues are summarised as follows:

- PAD evaluations and interlaboratory comparison show varied PAD results due to factors such as:
  - types and degree of variety in the attack methods assessed,
  - the materials used for making artefact and the moulds for developing the artefacts,
  - variations in applying the attack methods, for example of different thicknesses,
  - the qualities of artefacts replicated,
  - the level of skills of the tester and the staff replicating the attack
  - environmental condition of the testing laboratory
- No common adversarial/penetration test methodology exists to determine presentation attack resistance. Precise descriptions of the presentation attack methods to be tried are

not openly published because there is a security risk in making them available. Standards for adversarial testing are under development.

- For consistency and repeatability, a number of required factors such as expert judgement, background knowledge, and skills vary between evaluators and test laboratory experts thereby consistency cannot be guaranteed.
- Interlaboratory comparisons help identify the factors most critical for repeatable performance.
  - Round robin evaluations of a particular TOE passed between organizations, and comparing results
  - Sharing of information regarding attack methods and sharing of specific artefacts
  - Mutual audit between test organizations of the methodologies used, and mutual training.
- Sharing of artefacts between the test organisations improves the repeatability and agreement is needed on the
  - Selection of attack types against which the TOE is potentially vulnerable, • the grading of the attack success rates of the performed attack types
  - Assessment of attack potential of the performed attack types.
- Performing the confidential test methods exposes information that is confidential to the TOE supplier and evaluation client. In order to minimize evaluation costs, previous evaluation results can be reused. The CC sets down guidelines for reuse of items of any TOE against any CC functional and assurance requirements. Some schemes owners do not permit reuse of previous evaluation.
- Repeatability tests in the case study between test organizations focussed on the issues and factors indicating that good repeatability with some systems and the spread of results in others could be attributed to the system themselves rather than on the variability in the methodology or artefacts used.
- In CC evaluations, the CC certification authority monitors the evaluations conducted by the evaluators. However, there are no documentation to enable interchange of knowledge between evaluators. Although repeatability of results is an objective of the CC, the issue is that expert judgement, background knowledge, and skills required vary between evaluators and test laboratory experts thereby consistency cannot be guaranteed.
- Methods for identification of vulnerabilities of products include: public domain searches using search engines, vulnerability database check and technology search.

Issues and findings from evaluation and certification of biometrics for secure access control.

This topic considered the certification of access control systems against draft technical specification from relevant biometric standards.

- National scheme requirements seem to differ and no harmonisation exists. Harmonization of requirements for multiple applications depend on performance under different schemes, vary on the stricter requirements (maybe confidential), resistance to presentation attacks and tamper attacks. It also depends on the performance requirements to be assessed via a scenario or a technology evaluation.
- Use of detection error trade-off (DET) or receiver operating characteristics (ROC) curves - in general, biometric access control product, comprising both biometric sensor and comparison algorithms operate at a fixed threshold calibrated by the supplier to attain a desired false match rate requirement. However, the performance algorithm can be evaluated over a range of thresholds depending on the use-case scenario.
- Metrics for transaction times are used to minimize the effect of outlier long transaction durations due to poor habituation or interruptions during the evaluation process and differ between test organizations due to different methodologies.
- Limitations of repeatability of scenario based performance evaluations are well understood in order to achieve consistency of results. When comparing products, similar devices based on the same biometric sensor model but tested in different years, showed quite consistent GFRR performance when tested under the same general conditions.
- Certificate validity & Periodic assessment period of two years is appropriate and suppliers are expected to support the certified version of the product for the duration of the certificate validity. It seemed appropriate for the access control use case and in keeping up with the product updates.
- Standards covering assessment of presentation attack resistance, are under development. Evaluations of attack resistance are often inconsistent between test organisations, mainly due to variations in expertise, materials, artefacts, costs and resources.
- Determination of tamper resistance - Resistance to tamper attack is assessed by evaluating methods by which the attacker could gain access to the internal components of the device without causing a tamper alarm on the system.
- Biometric detection error rates are statistical measures, based either on

  (i) biometric samples collected from test subjects via the biometric product in scenario evaluation,

  (ii) a previously collected dataset of suitable biometric images in a technology evaluation.

  In both cases the measured error rates will be imprecise due to statistical error, and due to differences between the biometric samples used in the evaluation and those that would occur in the target application. Statistical measurement should use a sufficient number of test subjects and test samples to deliver required statistical significance levels, and statistical significance should be assessed.