Performance and Vulnerability Testing according ISO/IEC: overview and relevance

Christoph Busch

Hochschule Darmstadt / European Association for Biometrics http://www.christoph-busch.de

> EAB Members Meeting Frankfurt November 17, 2014









Agenda

- Certification is important and requested
 - should one develop new standards to test against?
 - should one test against existing standards?
- Existing International Standardisation
- Performance Testing
- Privacy Compliance Testing
- Vulnerability of Biometric Capture Devices
 - Development of the Presentation Attack Detection Standard

International Standardisation

Biometric Standardisation



Christoph Busch

ISO/IEC SC37 Biometrics

Established by JTC 1 in June 2002 to ensure

• a high-priority, focused and comprehensive approach worldwide for the rapid development of formal generic biometric standards

Scope of SC37

- "Standardization of generic biometric technologies pertaining to human beings to support interoperability and data interchange among applications and systems. Generic human biometric standards include: common file frameworks; biometric application programming interfaces; biometric data interchange formats; related biometric profiles; application of evaluation criteria to biometric technologies; methodologies for performance testing and reporting and cross jurisdictional and societal aspects"
- http://www.jtc1.org

Next meeting: January, 2015

Working Group 3

Title: Biometric Data Interchange

- Convenor: Christoph Busch (Germany)
- Terms of Reference:
 - To consider the standardisation of the content, meaning, and representation of biometric data formats which are specific to a particular biometric technology. To ensure a common look and feel for Biometric Data Structure standards, with notation and transfer formats that provide platform independence and separation of transfer syntax from content definition

"Getting equipment to understand each other"

Working Group 5

Title: Biometric Testing and Reporting

• Convenor: Nigel Gordon (UK)

Terms of Reference:

- To create testing and reporting methodologies and metrics that cover biometric technologies, systems and components
- To develop Working Drafts for approved projects on biometric testing and reporting.

"how to check it works"

Biometric Standardisation

Onion Layers

- Layer 1: BDIR
 - Digital representations of biometric characteristics
- Layer 2: LDS
 - CBEFF Meta-data
- Layer 3+4: System properties
 - Security
 - Performance
- Layer 5: BioAPI, BIP
 - System Integration



Levels of Development?

Progression levels

- Working Draft (WD)
- Committee Draft (CD)
- Darft International Standard (DIS)
- Final Draft International Standard (FDIS)
- International Standard (IS)

Issues to consider:

- Need for mature technology
- Decisions are made on consensus
- Commenting periods
- Potentially multiple loops at one level
- Need to progress
- Five year revision cycle



ISO/IEC Interchange Format Standards



The 19794-Family: Biometric data interchange formats

Generation 2 of ISO/IEC 19794



the semantic is equivalent for binary encoded and XML encoded records

		D	
Christ	ionn	INNSC	n
	opri		

Part 6: Iris image data

ISO/IEC 19794-6:2011



- 4 new iris image formats, compressible to as little as 2,000 bytes
- Iris formats are now highly empirically based, thanks to NIST IREX testing results
- Recommended target record sizes for different applications
- Recommended compression for different applications
- Formats differ in their required amount of image pre-processing
- Original 19794-6:2005 raw image format retained as one case
- Iris sample quality (29794-6) will become normative Annex

Biometric Performance Testing ISO/IEC 19795-1:2006

Biometric Performance Testing

Operators may think:



Biometric Performance Testing Standard

ISO/IEC 19795-x, Information technology -Biometric performance testing and reporting

- Part 1: Principles & Framework
 - Guidance applicable to the broad range of tests
- Part 2: Testing Methodologies for Technology and Scenario Evaluation
 - Multiple visits, habituation, enrolment
- Part 3: Modality-Specific Testing
 - Modality (& application) specific methodologies
- Part 4: Interoperability Performance Testing
 - Performance on other vendors data
- Part 5: Framework for biometric device performance evaluation for access control
- Part 6: Testing Methodologies for Operational Evaluation
- Part 7: Testing of ISO/IEC 7816-based Verification Algorithms
- Part 8: Performance Testing of Template Protection Schemes

Probability densitiv Distribution Function (PDF) $\Phi_g(s)$: PDF of genuine similarity score s(Q, R) $\Phi_i(s)$: PDF of imposter similarity score s(Q, R)

False-Match-Rate (FMR)

- **Def in ISO-HBV**: proportion of the completed biometric non-mated comparison trials that result in a false match
- Note: non-mated comparison trials are also referred to as impostor trials

$$FMR(t) = \int_{t}^{1} \Phi_{i}(s) ds$$



Probability densitiv Distribution Function (PDF) $\Phi_g(s)$: PDF of genuine similarity score s(Q, R) $\Phi_i(s)$: PDF of imposter similarity score s(Q, R)

False-Non-Match-Rate (FNMR)

- **Def in ISO-HBV**: proportion of the completed biometric mated comparison trials that result in a false non-match
- Note: mated comparison trials are also referred to as genuine trials

$$FNMR(t) = \int_0^t \Phi_g(s) ds$$



Algorithm error rates

- False-Match-Rate (FMR) often confused with FAR
- False-Non-Match-Rate (FNMR) often confused with FRR
- Equal-Error-Rate (EER)



Categorization

- Technology testing
 - Algorithmic level verfication error
 - False-Match-Rate (FMR) algorithm accepts "zero-effort" imposter
 - False-Non-Match-Rate (FNMR) algorithm rejects true identity
- Scenario testing and operartional testing
 - System level verification error
 - False-Accept-Rate (FAR)
 - False-Reject-Rate (FRR)
 - System level error requires observation of:
 - Sample generation: Failure-to-Capture (FTC)
 - Enrolment: Failure-to-Enrol (FTE) no reference for this subject
 - Verification: Failure-to-Acquire (FTA) no probe feature vector

False-Accept-Rate (FAR)

$$FAR = FMR * (1 - FTA)$$

False-Reject-Rate (FRR)

$$FRR = FTA + FNMR * (1 - FTA)$$

Failure-to-Capture Rate (FTC)

aligned to ISO-HBV: proportion of failures of the biometric capture process to produce a captured biometric sample

$$FTC = \frac{N_{tca} + N_{nsq}}{N_{tot}}$$

where

 $\begin{array}{ll} N_{tca} & \text{is the number of terminated capture attempts} \\ N_{nsq} & \text{is the number of images created with insufficient sample quality} \\ N_{tot} & \text{is the total number of capture attempts} \end{array}$

Failure-to-eXtract Rate (FTX)

Def: proportion of failures of the feature extraction process to generate a template from the captured biometric sample

$$FTX = \frac{N_{ngt}}{N_{sub}}$$

where

 N_{ngt} is the number of cases, where no template was generated N_{sub} is the total number of biometric samples being submitted to the feature extraction

Failure-to-Enrol Rate (FTE)

aligned to ISO-HBV: proportion of a specified set of biometric enrolment transactions that resulted in a failure to create and store a biometric enrollment data record

$$FTE = \frac{N_{nec}}{N}$$

where

- N_{nec} is the number of cases, where the biometric characteristic of the subject can not be captured at all.
- N is the total number of subjects, intended to be enroled in the biometric application

Failure-to-Acquire Rate (FTA)

aligned to ISO-HBV: proportion of a specified set of biometric acquisition processes that were failure to accept for subsequent comparison the output of a data capture process

Note: This is caused by either a FTC or an FTX in the in the verification process. No probe feature vector.

$$FTA = FTC + FTX * (1 - FTC)$$

Graphical Presentation



Graphical Presentation

ROC curve (Receiver operating characteristic curve)

 Plot of the rate of false positives (i.e. impostor attempts accepted) on the x-axis against the corresponding rate of true positives (i.e genuine attempts accepted) on the y-axis plotted parametrically as a function of the decision threshold



Graphical Presentation

DET curve (detection error trade-off curve)

 modified ROC curve which plots error rates on both axes (false positives on the x-axis and false negatives on the y-axis)



Testing Privacy Compliance



Operators may think:



Revocability ?

Data subjects may think:

"The number of biometric characteristics is limited (e.g. we have only 10 fingers) we can not revoke the biometric reference"



Data Protection Requirements

Requirements for data privacy and data protection are formulated in:

- Directive 95/46/EC: On the protection of individuals with regard to the processing of personal data and on the free movement of such data
- EU data protection regulation under development since 2012 http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf
- Regulation 45/2001: on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:008:0001:0022:en:PDF
- Directive 2002/58/EC: concerning the processing of personal data and the protection of privacy in the electronic communications sector

http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:FIN:EN:PDF

Biometric Template Protection

We do NOT store fingerpint, iris or face images

- we transform templates to pseudonymous identifiers (PI)
- we reach
 - Secrecy: biometric references (PI) can be compared without decryption.
 - Diversifiability / Unlinkability: Unique pseudonymous identifier can be created for each application to prevent database cross-comparison
 - Renewability: we can revoke and renew template data.
 - Noise-robustness: Stored information can be used for authentication with noisy biometric samples
 - Non-invertibility: Original biometric sample can not be reconstructed

[Br2008] J. Breebaart, C. Busch, J. Grave, E. Kindt: "A Reference Architecture for Biometric Template Protection based on Pseudo Identities", in BIOSIG-2008, GI-LNI, (2008) http://www.christoph-busch.de/files/Breebaart-BTPReferenceArchitecture-BIOSIG-2008.pdf

Biometric Template Protection

Protection at the same accuracy level is possible

Bloom filter-based pseudonymous identifiers



Data Protection Requirements

A technical guideline, how to implement requirements for data privacy and data protection is formulated in:

• ISO/IEC 24745: Biometric Information Protection, (2011) http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=52946



ISO/IEC 24745 Biometric Information Protection !

Vulnerability Testing

Vulnerability Testing

Operators may think:

"Biometric sensors can not detect gummy and cut-off fingers"


Presentation Attacks



Gummy Finger Production in 2000 !

Attack without support of an enroled individual

- Recording of an analog fingerprint from flat surface material
 - z.B. glass, CD-cover, etc. with iron powder and tape
- Scanning and post processing:
 - Correction of scanning errors
 - Closing of ridge lines (as needed)
 - Image inversion
- Print on transparent slide
- Photochemical production of a platine



Gummy Finger Production in 2000 !

Reported in a publication by BKA

 A. Zwiesele et al. "BioIS Study - Comparative Study of Biometric Identification Systems", In: 34th Annual 2000 IEEE International Carnahan Conference on Security Technology, Ottawa, pp. 60-63, (2000)

BioIS Study

Comparative Study of Biometric Identification Systems

4.) Dupability: The aim of this part is to analyse and assess the effort that is necessary to dape biometric systems. It not only covers the systems taking part in the issue of the systems taking part in the issue of the system taking part is the system of the system of the system implementation.

implementation.
5.) Influence of the various programmable system parameters: This part attempts to investigate the repercussions of the various system setups for the identification attributes. The findings are intended to

permit recommendations to be made reparding the prefered settings for each of the biometric systems under investigation. 6.) Influence of the various environmental factors on the identification reliability of the systems: The purpose of this part is to determine the

repercussions of changes in environmental conditions for the identification attributes. One example of such factors might be the

audience of the results of the study and the

"In comparison to PINs and passwords, a biometric signature has crucial advantages and provides an unambiguous proof of identity..." "Comprehensive empirical tests are being conducted to get rid of the last doubts and insecurities from the angle of consumer and data

",Widespread employment of biometric systems ju: around the corner..."

...that is what the manufacturers are promising, but as a study by the Federal Criminal Investigation

knowledge which could be gained.

Introduction

A.Zwiesele, BKA Wiesbaden³ - A.Munde, BSI Bonn³ Dr. C.Busch, H.Daum, IGD Darmstadt³

Abstract

On 1st April 1999, after a preparatory phase lasting more than twelve months, work on the a.m. BiolS Study finally commenced. This study was similated by the Federal Criminal Investigation Office of Germany (IRA) in close cooperation with the German Information Security Agency (IBS). The study was executed by the Fraunhofer Institute of Graphical Data Processing (IGD).

The study includes a field investigation, in which 11 physiological (static) and behaviour-specific (dynamic) systems, which were available and supported in Germany, were installed and put into operation in a defined senarion. The field investigation was conducted with approximately 40 users representing different age, employment, educational and ethnic groups.

The main objectives of the field investigation are as follows:

 To gather experience with the biometric systems and to identify any weaknesses that need to be examined in greater depth during the future course of the study.
 To obtain statistical information regarding the frequency with which authorised users are rejected by the various systems. This

 information will then be taken as a basis for establishing the existence of certain user groups which individual systems have difficulties in identifying. In the event that such groups do exist, the possible reasons for their rejection need to be examined.
 To observe the behaviour of the users over a prolonged period of time, in order to

a prolonged period of time, in order to establish whether or not any changes can be observed. There might, for instance, be a certain familiarisation effect, which is reflected in a change in the rejection rate.

The field investigation is to be followed by a further technical study phase, designed to investigate the following points:

Federal Criminal Investigation Office of Germany
 ² German Information Security Agency
 ³ Fraunhofer Institute of Graphical Data Processing

0-7803-5965-8/00/\$10.00 @2000 IEEE

Vulnerability Testing

2014-11-17

Face Mask Production in 2013

Attack again without support of an enroled individual

- Frontal and profile photos are uploaded
- 3D face dataset rendered and produced





Capture process

• Camera operating in macro modus



Preview image of the camera with LED on (left) and LED off (right)

LED permanent on



Finger illuminated

[SNB12] C. Stein, C. Nickel, C. Busch, "Fingerphoto Recognition with Smartphone Cameras", Proceedings 11th Intern. Conference of the Biometrics Special Interest Group (BIOSIG 2012)

Finger recognition study - 2012/2013

• Results: biometric performance at 1.2% EER



DET Curve

[SBB13] C. Stein, V. Bouatou, C. Busch, "Video-based Fingerphoto Recognition with Anti-spoofing Techniques with Smartphone Cameras", Proceedings 12th Intern. Conference of the Biometrics Special Interest Group (BIOSIG 2013)

Finger recognition study - 2012/2013

- Presentation Attacks
 - 1: replay from Smartphone display (simple)
 - 2: self generated print-outs (not critical to detect)
 - 3: Ralph Breithaupt's / BSI best artefacts (very challenging)



Replay attack



Simple artefacts



Challenging artefacts

Finger recognition study - 2012/2013

- Observation
 - significant strong light reflection near the fingertip
 - from the cameras LED
- Reflection depends on
 - Shape of the finger
 - Consistency of the finger
 - Angle of the finger to the camera
- Attack detection, as light reflection differs from artefacts to genuine fingers



 [SBB13] C. Stein, V. Bouatou, C. Busch, "Video-based Fingerphoto Recognition with Anti-spoofing Techniques with Smartphone Cameras", Proceedings12th Intern. Conference of the Biometrics Special Interest Group (BIOSIG 2013)

Finger recognition study - 2012/2013

• Normal presentation with genuine finger

Genuine Finger (Challenge Response)



Results: Presentation Attack Detection (PAD)



Conclusion:

better Presentation Attack Detection than capacitive sensors

Liveness Detection

ISO/IEC CD 30107 - Presentation Attack Detection

Attacks on Biometric Systems



Source: ISO/IEC 30107-1 inspired by N.K. Ratha, J.H. Connell, R.M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," IBM Systems Journal, Vol 40. NO 3, 2001.

ISO/IEC 30107 - Scope

- terms and definitions that are useful in the specification, characterization and evaluation of presentation attack detection methods;
- a common data format for conveying the type of approach used and the assessment of presentation attack in data formats;
- principles and methods for performance assessment of presentation attack detection algorithms or mechanisms; and

Outside the scope are

- standardization of specific PAD detection methods;
- detailed information about countermeasures (i.e. anti-spoofing techniques), algorithms, or sensors;
- overall system-level security or vulnerability assessment.

30107 parts

- Part 1 (IS) Framework
 - Elaine Newton
 - status: 2nd CD
- Part 2 (IS) Data formats
 - Olaf Henniger
 - status: WD
- Part 3 (IS) Testing and Reporting
 - Michael Thieme
 - status: WD

Definitions in ISO/IEC 30107 PAD - Part 1: Framework

• presentation attack

presentation to the biometric capture subsystem with the goal of interfering with the operation of the biometric system

presentation attack detection (PAD)

automated determination of a presentation attack

Definitions in ISO/IEC 2382-37: Vocabulary http://www.christoph-busch.de/standards.html

impostor

subversive biometric capture subject who attempts to being matched to someone else's biometric reference

• identity concealer

subversive biometric capture subject who attempts to avoid being matched to their own biometric reference

ISO/IEC 30107 -Examples of Artificial and Human Attack Presentation

Artificial	Complete	gummy finger, video of face
	Partial	glue on finger, sunglasses, artificial/patterned contact lens, non-permanent make up
Human	Lifeless	cadaver part, severed finger/hand
	Altered	mutilation, surgical switching of fingerprints between hands and/or toes
	Non-Conformant	facial expression/extreme, tip or side of finger
	Coerced ¹	unconscious, under duress
	Conformant	zero effort impostor attempt

Source: ISO/IEC 30107

ISO/IEC 30107 - Definitions

presentation attack instrument (PAI)

biometric characteristic or object used in a presentation attack

artefact

artificial object or representation presenting a copy of biometric characteristics or synthetic biometric patterns

Types of presentation attacks



Biometric framework with PAD



Source: ISO/IEC 30107

Reporting about the PAD using ISO/IEC 30107-3

PAD-Standard

Methodology in ISO/IEC 30107 Presentation Attack Detection - Part 3: Testing and reporting

- Security Evaluation
 - for evaluations using the Common Criteria Framework
 - Protection Profile (PP) (e.g. from German BSI)
 - Security Target (ST)
 - Evaluation Assurance Level (EAL)
 - Assessment of the attack potential
 - "if there is at least one aretefact that can reproducibly successful attack the PAD-component then the PAD failed the test"
- Other approaches
 - for evaluations in academic and technology development
 - tolerating the limited statistical significance of small test set
 - the statistical distribution is unknown and for sure not normal
 - " a score based metric can tell us, if the method improved"

PAD-Standard

Metrics in ISO/IEC 30107 Presentation Attack Detection

- Part 3: Testing and reporting
 - Attack presentation classification error rate (APCER) proportion of attack presentations incorrectly classified as normal presentations at the component level in a specific scenario
 - Normal presentation classification error rate (NPCER) proportion of normal presentations incorrectly classified as attack presentations at the component level in a specific scenario

Applying ISO/IEC 30107-3 Metrics

Do the metrics currently in ISO/IEC 30107 PAD - Part 3: serve to provide a meaningful report?

- [SBB13] Publication:
 - The reported number of attack presentations incorrectly classified as normal presentations was one out of four artefacts
- Thus the APCER to be reported is

$$APCER = \frac{1}{4} = 0.25$$

 but there were in fact 27 artefact species, that were used in the background but not reported as they are not challenging

$$APCER = \frac{1}{27} = 0.04$$

Thoughts for improving ISO/IEC WD 30107

Trust in a biometric sensor relates to risk

Apply classical risk assessment ?

- *Risk = Impact of Risk event x Probability of Occurrence*
- we do not know the impact!

Modified assessment

• *Vulnerability* = *Attack Potential* x *Probability of Occurrence*

Needed Change

- The size of the corpus with the artefact species is essential
- The CC-related attack potential should be included in the metric definition for non-cc evaluations
 - 30107-1: **attack potential** attribute of a biometric presentation attack expressing the effort expended in the preparation and execution of the attack in terms of elapsed time, expertise, knowledge about the capture device being attacked, window of opportunity and equipment, graded as "no rating", "minimal", "basic", "enhanced-basic," "moderate" or "high.
- The known success rate of a presentation artefact instrument is relevant and might be an approximation for the probability of occurrence

Suggested augmented metric for ISO/IEC 30107-3

- Attack presentation classification error rate (APCER) proportion of attack presentations incorrectly classified as normal presentations at the component level a specific scenario taking the attack potential and the known attack instrument success rate into account.
- Attack potential (AP) = {0.2 for "minimal", 0.4 for "basic", 0.6 for "enhanced-basic," 0.8 for "moderate", 1.0 for "high.}
- Presentation attack instrument success rate (PAISR)
 Proportion of evaluated capture devices that could be spoofed by the specific PAI (i.e. artefact).
 - would start with a value of 1 for a new discovered artefact species and could be reduced over time (as more sensors become robust)

Suggested refined metrics for ISO/IEC 30107-3

• The APCER could thus be expressed as

$$APCER = \frac{\sum_{i=1}^{N_{AS}} (RES_i * AP_i * PAISR_i)}{N_{AS}}$$

- N_{AS} number of presentation attack instruments (PAI) (i.e. artefact species) in the corpus
- RES_i result of attack with ith PAI {0 for detected attack, 1 for successful attack}
- AP_i attack potential of the ith PAI (close to zero, if artefact is easy to produce)
- $PAISR_i$ presentation attack instrument success rate (close to zero, if all sensor can detect this artefact)

Suggested refined metrics for ISO/IEC 30107-3

- Normal presentation classification error rate (NPCER): proportion of normal presentations incorrectly classified as attack presentations at the component level in a specific scenario
- The NPCER could thus be expressed as

$$NPCER = \frac{\sum_{i=1}^{N_{GPA}} RES_i}{N_{GPA}}$$

- N_{GPA} number of normal presentations from a genuine subject
- RES_i result of presentation detection component for the ith attempt {0 for no detected attack, 1 for false alarm}

Conclusion

The standardisation process is open process

- Register and contribute to ISO/IEC 30107
 Presentation Attack Detection
- Open question:
 - should PAD metrics and performance metrics be merged ?



Recent research provides effective countermeasures to detect artefacts

Vein recognition



Fingervein image

 Fingerprint Recognition with Optical Coherence Tomography (OCT)



3D Finger OCT scan

References

Web

- WG3 onvenors website with latest news http://www.christoph-busch.de/standards-sc37wg3.html
- ISO/IEC JTC SC37 http://isotc.iso.org/livelink/livelink? func=II&objId=2262372&objAction=browse&sort=name
- Wikipedia http://en.wikipedia.org/wiki/ISO/IEC_JTC_1/SC_37
- Published ISO Standards http://www.iso.org/iso/iso_catalogue/catalogue_tc/ catalogue_tc_browse.htm?commid=313770&published=on
- ISO/IEC 19795-1:2006, Biometric performance testing and reporting - Part 1:Principles and framework
- ISO/IEC 24745:2011, Biometric information protection
- ISO/IEC CD 30107-1, Biometric presentation attack detection
 Part 1: Framework
- ISO/IEC WD 30107-3, Biometric presentation attack detection
 Part 3: Testing and reporting

Visit Norway in 2015

Norwegian Biometrics Laboratory Workshop 2015

- Presentation Attack Detection in Biometrics: Solved and Unsolved Challenges
- Chair: Dr. Raghavendra Ramachandra
- Monday, March 2, 2015
- please follow us at: http://nislab.no/biometrics_lab

ISO/IEC JTC1 SC37 Conference

- Working Group Meetings
- June 22 to 26, 2015 in GUC
- Standards Norge
- We are seeking **Sponsors** for the ISO conference

Contact



Contact





