Outline of Presentation

- Scope
- Background
- QPL Process
- Contrasting Evaluation Reporting Methods
- Lessons Learned
Narrowing the Scope

Scope is Biometric Product Performance testing using Scenario evaluation techniques focused on Access Control applications using 1:1 Verification matching.
Background

Specific information relates to US DHS TSA QPL

- Mandated in 2004 (US Public Law 108-458)
- Requirements and Process published - 2005
- Initial QPL – 2007
- Addition to QPL – 2009
- Process in place, on-going
- Specific to airport PACS

DHS – Department of Homeland Security
TSA – Transportation Security Administration
QPL – Qualified Products List
PACS – Physical Access Control System
Biometric portion of PACS
QPL Process – Event Sequence

• Airports demand for biometric product for PACS
• Manufacturers submit to third-party testing (fee based)
• Suitable testing organization delivers performance report
• TSA approves product → QPL
• Airports select from QPL for purchase
Key Test Characteristics

- Test crew size >250 (enroll), >200 (verify)
  - Plan for “drop-out”

- 4 visits: Enroll, 3 Revisits (~2 weeks apart)
  - Enrollment includes training and initial verification

- 5 genuine & 5 imposter transactions/revisit
  - Minimum 3000 transaction (each type)
  - Imposters are randomized from other crew members
Performance Metrics and Requirements Definition

Pass/Fail testing must establish metrics, requirements and decision criteria:

- False Accept Rate $< 1\%$
- False Reject Rate $< 1\%$
- Transaction Time $< 6$ seconds
- Failure to Enroll Rate $< 3\%$

NOTES:
- Matching errors based on up to 3-attempt transactions
- Confidence bounds used to evaluate Pass/Fail
- Requirements set with industry, airport and NIST participation
Use of Confidence Bounds

Measured Error Rate with Confidence Intervals (arrows)

FAR

Required

Measured
Scenario Testing Approaches

- **Conformance to Performance Specification Testing**
  - E.g. TSA QPL
  - Report only “Pass or Fail” (as one operating point)

- **Graded Performance Testing**
  - ISO Standard 19795-5 or INCITS 409.5
  - Matching grades range 0 (worst) to 6

- **Pure Performance Testing**
  - E.g. NPL or IBG (CBP series)
  - DET or ROC (typically)
NVLAP Handbook 150-25, Figure A.3. Scopes of accreditation
Graded Performance

From ISO FCD 19795-5
• Shows 3-FAR level range (0.1 – 1.0%)
• Shows grades assigned at 6 FRR thresholds
  <1% = grade 4
Pure Performance Reporting

Transactional DETs (Same-Day and Different-Day)

From IBG CBT 6 public report
Contrast in Results Reporting

Pure Performance – maximum information, more difficult to interpret
Graded Performance – typical operating range, shows performance levels
Pass/Fail – simplest, no indication of preference

### Graded

<table>
<thead>
<tr>
<th>Target FAR</th>
<th>0.1%</th>
<th>0.3%</th>
<th>1.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRR Grade – BLUE</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>FRR Grade – RED</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>FRR Grade - GREEN</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

### Pure

Transaction DETs (Same-Day and Different-Day)

### Pass/Fail

<table>
<thead>
<tr>
<th>FRR Decision – BLUE</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRR Decision – RED</td>
<td>Pass*</td>
</tr>
<tr>
<td>FRR Decision – GREEN</td>
<td>Fail</td>
</tr>
</tbody>
</table>

* (depends on confidence level)
Lessons Learned

• Enrollment with multi-instance modalities
  • more complicated, needs written policy

• Identify need for re-enrollment
  • remove “avoidable goats”

• Criticality of threshold setting
  • responsibility – defined by process, enforced by lab, depends on supplier
Lessons Learned - 2

- Pass/Fail testing does NOT require access to matching scores (COTS configuration)
- Who controls Final Report dissemination to qualifying organization (lab or supplier?)
- Cost to participate is a deterrent to suppliers
Thanks for your attention

Contact information:

Rick Lazarick, CBP
CSC – Chief Scientist, Identity Labs
rlazarick@csc.com
609-883-6767