Biometrics - Operational Testing Experience in Airport Security

Biometrics Consortium 2004
September 20, 2004

Transportation Security Administration

Presentation Overview

- Introduction to NSSA
- Recent testing activity and results
- Reflection on testing experience – lessons learned

Role of National Safe Skies Alliance (NSSA)

- NSSA is a 501c3 non-profit consortium- Knoxville, TN (headquarters)
- Funded through a cooperative agreement with the TSA
- Dedicated to testing Aviation Security Technology (biometrics ~ 25% of total testing)

Knoxville Biometrics Test Door

- Used for physical access control biometric evaluations
- 9 devices tested to date
  - Fingerprint (5)
  - Facial
  - Iris
  - Hand Geometry
  - Fingerprint & iris
- Devices tested consecutively for ~3 yrs.

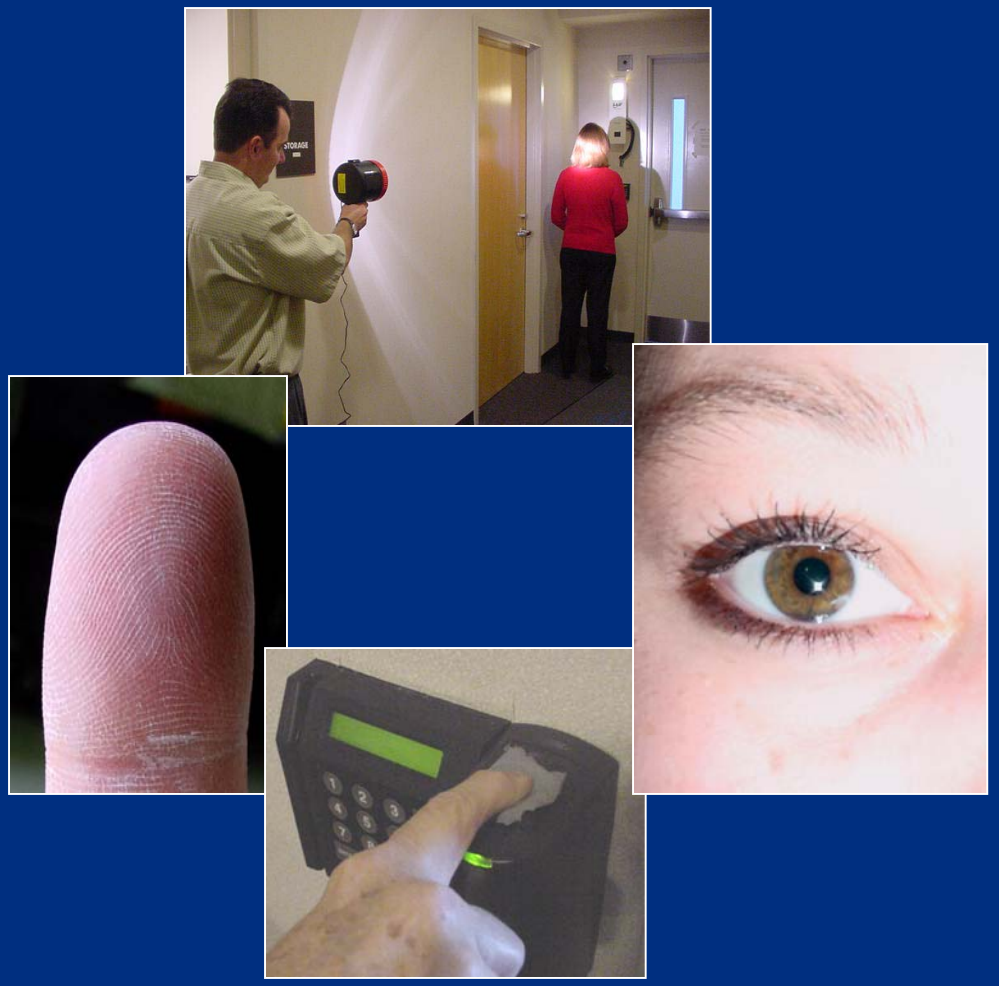
Transportation Security Administration

Other Biometric Related Testing

- Outdoor fingerprint – employee access control
- “Chokepoint” surveillance
- Wide area surveillance (in airport public area)

Typical Testing Scenarios

- Normal usage by employees
- Normal usage and “abnormal conditions” by test subjects
- False Accept by test subjects
- Attempts to defeat by test subjects



Transportation Security Administration

Recent Testing Results

4 Tests completed since last BC 2003 reporting

- Indoor fingerprint (CrossMatch)
- Outdoor fingerprint (Sagem)
- Indoor fingerprint and iris (NextGenID)
- Wide area surveillance (InteliTrac)

Results – Indoor Fingerprint

Employee Operational Data (only) - measured

FTE	0/27	Enrollment time	1:33
FNMR	19.2% (+/-8.1)	Transit time	0:08.1
	17.1% (+/-7.5)		0:10.5

Cumulative False Reject Rate = 3%

False accept * 0/557

* (not employees, but test staff acting as imposters)

Results – Outdoor Fingerprint

Employee Operational Data (only) - measured

FTE	1/86	Enrollment time	1:41
FNMR	5.8% (+/-1.9)	Transit time	0:09.4

Cumulative False Reject Rate = 3%

False accept *

* (assumed very low based on previous indoor evaluation of identical device)

Transportation Security Administration

Results – Indoor Fingerprint and Iris

Employee Operational Data (only) - measured

FTE	1/36	Enrollment time	4:10
Finger Only			
FNMR	26.7% (+/-7.5)	Transit time	0:14.2
Finger & Iris			
FNMR	29.5% (+/-8.0)	Transit time	0:15.3

Cumulative False Reject Rate = 4.1% Finger Only

8.8% Finger+Iris

False accept * 0/508

* (not employees, but test staff acting as imposters)

Transportation Security Administration

Results – Wide area surveillance

Note: Test subjects (seeds) used to establish ground truth

Subject presentations	376
Failure to Acquire	13%
Percent correct matches (acquired)	56%
Percent false matches (acquired)	24%

OT&E "Lessons Learned" Categories

- Test Coordination
- Test Planning
- Test Conduct
- Reporting of Data
- Statistical Methods

Lessons Learned-Test Coordination

- Matching a technology with a willing host is not as easy as it sounds
- The host organization should have a vested interest in the results
- The person assigned to work with you may not have the same interest level as the person who decided to host the test
- IMPORTANT-Installation should be where subjects have an **incentive for using the technology and do not have the opportunity to avoid using it**.

Lessons Learned- Test Coordination – Con't

- Identify access control integration issues early
  - Wide variation in access control systems
  - Many organizations no longer have support contracts in place
  - May need to consider a third-party integrator
- Need to approach suppliers with a previously developed plan
- Some suppliers genuinely want to learn from the evaluation; others simply want to use it as a *marketing tool*

Transportation Security Administration

Lessons Learned- Test Planning

A suggested testing methodology:

- Identify measurable parameters that can be used to judge the effectiveness of the technology
- Measure the parameters at the application site prior to technology insertion to provide baseline values
- Insert the technology and verify that it is operating per the manufacturer's specifications
- Measure parameters again after insertion of the technology into the operational environment
- Compare the baseline and "with technology" values to provide an evaluation of technology effectiveness

Lessons Learned- Test Planning – Con't

- Some flexibility is required in the test plan
  - Many factors are beyond the control of the tester
  - Virtually impossible to anticipate all scenarios
- **Match score values** may be hard to obtain
  - May only be practical to test at one threshold
  - Getting scores from device is problematic
- Plan on collecting more data points than required-some will be lost in the analysis
  - Plan to review data regularly
  - **Videotape is invaluable** for resolving issues

Lessons Learned- Test Planning – Con't

- Write detailed test procedures
  - Process provides a good simulation of actual testing; identifies problem areas
  - Provides a fill-in test coordinator with consistent process
- Build in fallback positions and an exit strategy into integration ("do no harm")
  - May need an easy way to turn off biometric under certain situations
  - Device may perform so poorly that it has to be removed prior to test completion

Lessons Learned- Test Conduct

- Best to have a single person oversee all testing
- Insure adequate training is provided to testers and users
- Allow for learning curve
  - 2 weeks seems about right
- **Be vigilant of behavior and performance by test subjects**
  - Intentional misuse or abuse
  - Gaming
  - Avoiding use by administrative means
  - Goats ("There's one in every crowd")

Lessons Learned- Test Conduct – Con't

- **Automate data collection** as much as possible
  - do not assume automated data collection system will tell the whole story
- Small incentives make a BIG difference in attracting short-term test subjects
  - Food vouchers work extremely well

Lessons Learned- Reporting of Data

- Note the version/model of the system evaluated
- Compare technology data against baseline values if available
- Look for demographic factors
- Look for relationships between enrollment quality and performance
- Use data to show operational impacts as well as technology performance
- Note maintenance, reliability, usability issues
- Detail how data was obtained and point out limitations of the data

Lessons Learned- Statistical Methods

- Challenging to balance time and cost constraints with sample size requirements
- Validate device data collection features before relying on them to provide test results
- Tests must be randomized to minimize potential system biases

Lessons Learned- Statistical Methods – Con't

- Because results are participant dependent, it is very important to consider demographics when selecting participants
  - Age
  - Gender
- Results are test subject dependent so normal usage results must be factored into abnormal condition results

Lessons Learned- Statistical Methods – Con't

- Non-parametric and parametric statistical techniques necessary to analyze results
  - Non-parametric techniques typically used with qualitative data (binomial distribution) such as FRR and FAR results
  - Parametric techniques typically used with quantitative data (normal distribution) such as enrollment and throughput times

Contact Information

Rick Lazarick

TSA Research and Development

Building 315, TSL-300

William J. Hughes Technical Center

Atlantic City International Airport, NJ 08405

609-813-2719 (voice)

609-383-1973 (fax)

Rick.lazarick@faa.gov

Richard.lazarick@dhs.gov

Transportation Security Administration

Transportation Security Administration