Biometric Testing And

Jean-Christophe.Fondeur@sagem.com

Sagem Sécurité
SAFRAN Group

Introduction

Biometric testing

- Objectives
- Test protocole
- Test database
- Test criteria

Performance extrapolation

- Accuracy
- Sizing

Conclusion

Testing Large Scale Biometric Systems

Operational system

- Large scale
- Complex

Observation: Test results

- Limited
- Potentially biased

We need:

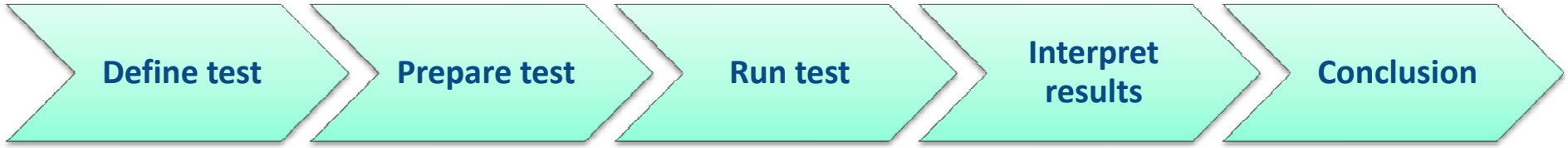
- **To characterize the target system**
- **A reliable observation**

  => Test protocole, database, criterias
- **A way to extrapolate from this observation**
  - Accuracy & Sizing
- **A risk analysis approach**
  - No testing is perfect …

Focus on

- **large scale back end systems (capture, human factors not in the scope)**
- **Biometric aspects (architecture, security not in the scope)**

Testing overview



- objectives

- Protocole
- Database
- Criteria

- Tuning
- Blind test

- Qualitative analyze
- Extrapolate results

- With regards to original objectives

Biometric Testing: Objectives

Testing can have many objectives

- **Check feasibility**
- **Estimate system performance, cost, risk**
  - Accuracy, HW sizing, operator workload, …
- **Select best technology provider**
- **Make some key system design choices**
  - Mono/multi modal, enrolment workflow,

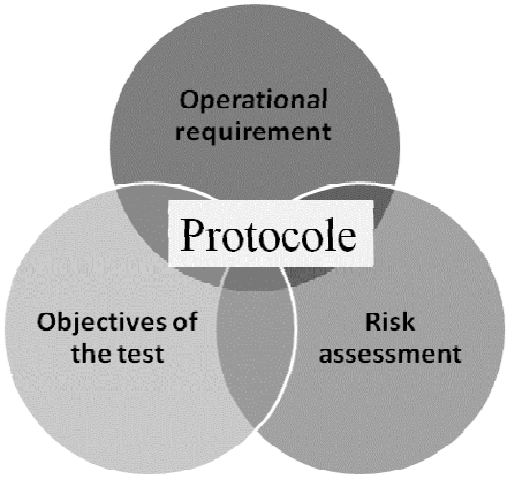
Test shall be explicitely designed for the desired purpose

- **Impact on test protocole, type and size of database, …**
- **Specific sub test may be required to reach some of the objectives**

Test Protocole

Test Protocoles are defined from

- **Objectives of the test**
  - « Why are we doing this test »
- **Operational requirement**
  - « What the system should do »
- **Risk assessment**
  - « What are the main risks ? »



Several subtest scenarii are often necessary

- **Accuracy, throughput, resistance to fraud & erors, …**

No test will ever be perfect, biases must be identified and analyzed.

Test preparation: Database

Test database must be

- **Representative of system scenario**
  - Acquisition conditions and workflow, Population characteristics, …
- **Unbiased**
  - Sequestered (blind test), not correlated with automated system
  - Mix of database shall be avoided (Better to conduct the test on each database)
  - Synthetic data are likely to introduce biases and shall be avoided
- **Large enough**
  - To capture diversity of situations
  - To enable extrapolation: database should be at least ~ 1/100 of the final system size
- **… and reasonably known and characterized** (Ground truth, quality distribution, …)

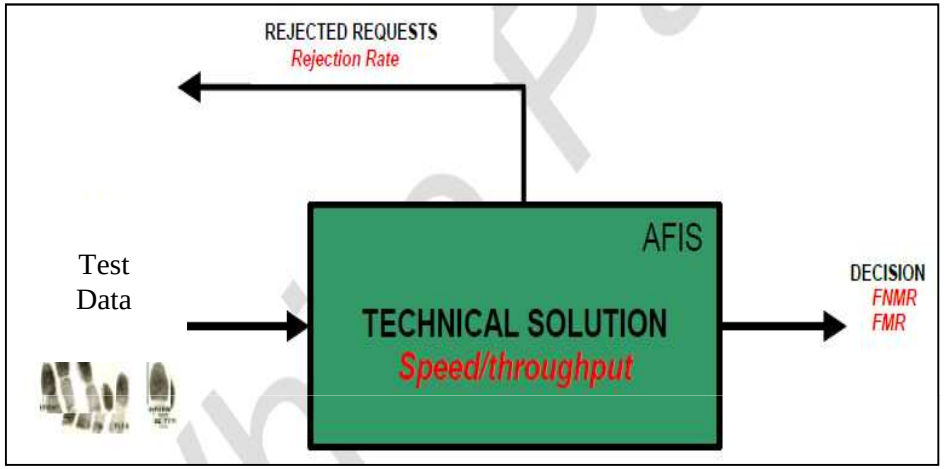
No database is perfect

- **Analysis shall be performed to assess risks and biases**

Test preparation: Criteria

Test criteria must be directly linked to system behavior

- Classical biometric indicators are
  - Rejection rate (FTE/FTA)
  - Accuracy (FMR, FNMR)
  - Throughput (Speed)



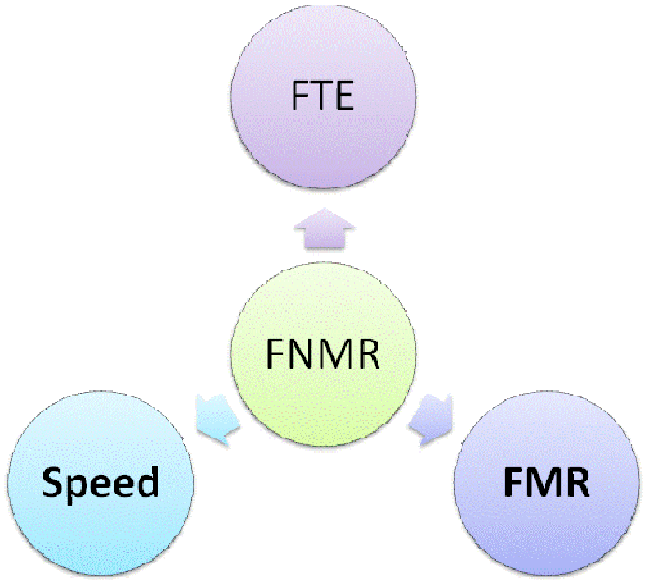
- Other indicators can be
  - Robustness to biometric errors & fraud
  - Interoperability, ….

- Internal parameter may be measured to help modeling system behavior
  - Filtering rate, number of correct minutiae, ….
  - They shall not be directly used as system performance criteria

- Specify the information that will be needed to calculate and interprete the results

Test preparation: Criteria

- Those Indicators are linked
  - FMR ⇔ FNMR: « *Decision policy* »
  - Speed ⇔ FNMR: « *Tuning policy* »
  - FTE/FTA ⇔ FNMR: « *Rejection policy* »



⇒ Those policies are business policies, not technology policies

- Weight of those criteria must be known before testing to enable system tuning

⇒ Tuning to system requirement is necessary

- Tuning to business policies and -to some extent- to system data

⇒ Those indicators must be measured simultaneously

Run test …

Set up
- Install, tuning, smoke test

Blind test
- no modification to system

Documentation, logging and monitoring are critical during this phase

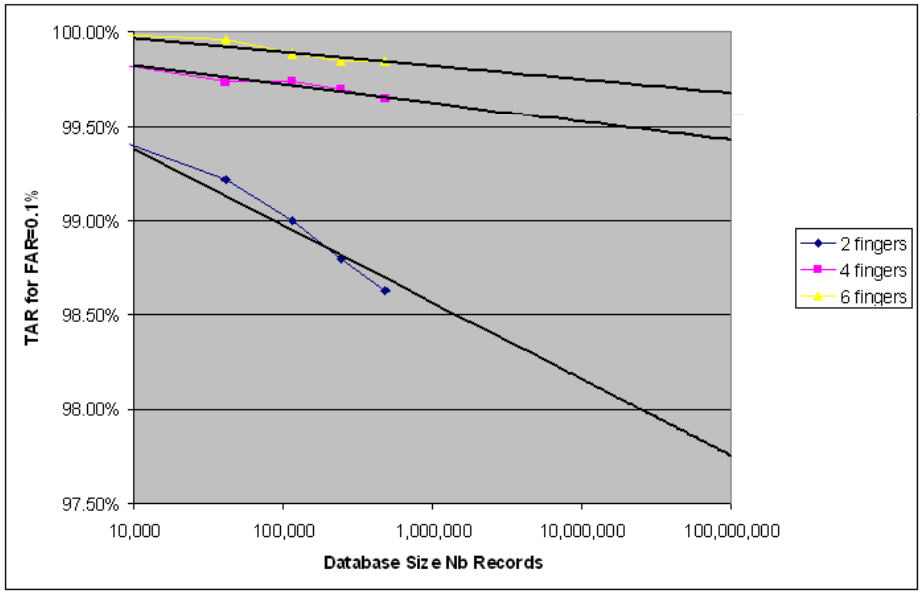
- Biometric results
- Timing information
- Errors and anomalies
- System behavior (CPU, I/O, …)

Interprete results

- **First thing to do is to check and validate the results**
  - **Calculate performance indicators**
  - **Is blind test consistent with the tuning tests ?**
    - Accuracy, speed
  - **Analyze potential errors (corrupted files, problems in the data, …)**
  - **Validate ground truth**
    - Are there unexpected hits ?
    - Are there any surprising results ?
- **Then you can do more interpretation**
  - **Detailled analysis**
    - Correlation with data characteristics
  - **Extrapolation**

Accuracy Extrapolation

- **Use statistical methods (parametric and non parametric):**
  - See “Biometrics system: Technology, Design and performance evaluation”, Springer, 2005 by Wayman, Jain, Maltoni, Maio, p 263-287
- **Use empirical methods**
  - Plot TAR vs Database size



- **Raises fundamental statistical issues**
  - Independence of measurements
- **In practice, Simplistic extrapolation approach**
  - **Provide acceptable results (**if extrapolating by less than 100)
  - **Permits to project conclusions** (at least qualitatively)

- ROC curves: FNMR (FMR, 10xDBSize) ~ FNMR (FMR/10, DBSize)

Sizing Extrapolation

- **Extrapolating sizing from test measurements is very complex**
  - **Biometric factors**
    - Algorithm speed (can be estimated in testing providing test DB is large enough)
  - **Non biometric factors**
    - Architecture considerations, hardware limitation, …

- **However simple « rule of three » on matching time can and must be done**
  - **It is certainly not sufficient to prove scalability**
  - **It is often sufficient to prove non-scalability**
    - Throughput = 100,000 requests/day
    - Database size =100,000,000 people
    - Measured matching speed = 1,000 record / sec / server

      => Requires over 115,000 servers …

Conclusion

- Testing is application dependent
  - Need to define objectives, protocole, database and criteria
- Main biometric criteria (FTE/FMR/FNMR/Speed) are linked
  - Trade off must be decided by business, not by technology
  - Some tuning (to policies and to data) is necessary
- There are biases in every test and extrapolation
  - They need to be minimized
  - They need to be known and taken into account when drawing conslusions
- Simplisitic extrapolation techniques are useful (even if not sufficient)

Thanks!