Latent Test Sets
"Wish List"

Vladimir Dvornychenko

Steve Wood & Michael Garris

April 6, 2006
NIST

Characteristics of Ideal Testing Set

- Statistically adequate number of samples

- Representative of community

Statistical Considerations

- The more samples the better – but…
- Reality dictates pick the least number adequate to estimate parameters of interest
- Execution time is also a concern
- Compromise ➜
  - ~ 1000 probes
  - ~ 10,000 background
  - Even this may be too computationally demanding

Test Set Should be Representative

- The test set should include samples from many sectors of the latent community
- And it should not be overrepresented by one sector
  - For example, using hits obtained by system "A" to draw samples from may result in a population biased toward system "A"
  - Picking samples "by hand" may introduce other biases
  - A multitude of sources and sampling approaches is desired

Data Ownership and Protection

- Always a concern!

- Significantly influences how tests may be conducted

- NIST treats fingerprint data as "Law Enforcement Sensitive," and we do not redistribute data

Federal Bureau of Investigation
www.fbi.gov

Homeland Security

Need Innovative Sources of Data

- Outreach to smaller agencies
  - 50 samples each from a large number of smaller agencies can lead to a broadly representative test set
  - This approach will provide more diversity/representation

- Larger agencies may need to go into their legacy files to provide samples not selected by an AFIS

Is there a Need for "Special Purpose" Test Sets?

- Test sets for identifying and quantifying weaknesses
- Primary purpose is not to rank performance of competing vendor systems
- May involve unusual data types and graduated exemplars of known difficulties
  - Same print with decreasing number of minutiae
  - Same print with increasing shear/distortion
  - Different methods for lifting print (ninhydrin, iodine, superglue, laser, various powders)
  - Unusual pattern classes
  - Scars
- Such data sets will probably involve human volunteers