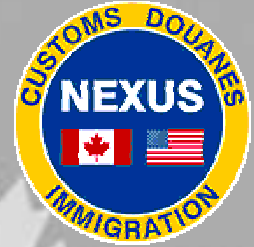




Canada Border
Services Agency

Agence des services
frontaliers du Canada



Score calibration for optimal biometric identification

(see also NIST IBPC 2010 online proceedings:
<http://biometrics.nist.gov/ibpc2010>)

AI/GI/CRV 2010, Ottawa

Dmitry O. Gorodnichy

Head of Video Surveillance and Biometrics Section (2008-)

Richard Hoshino

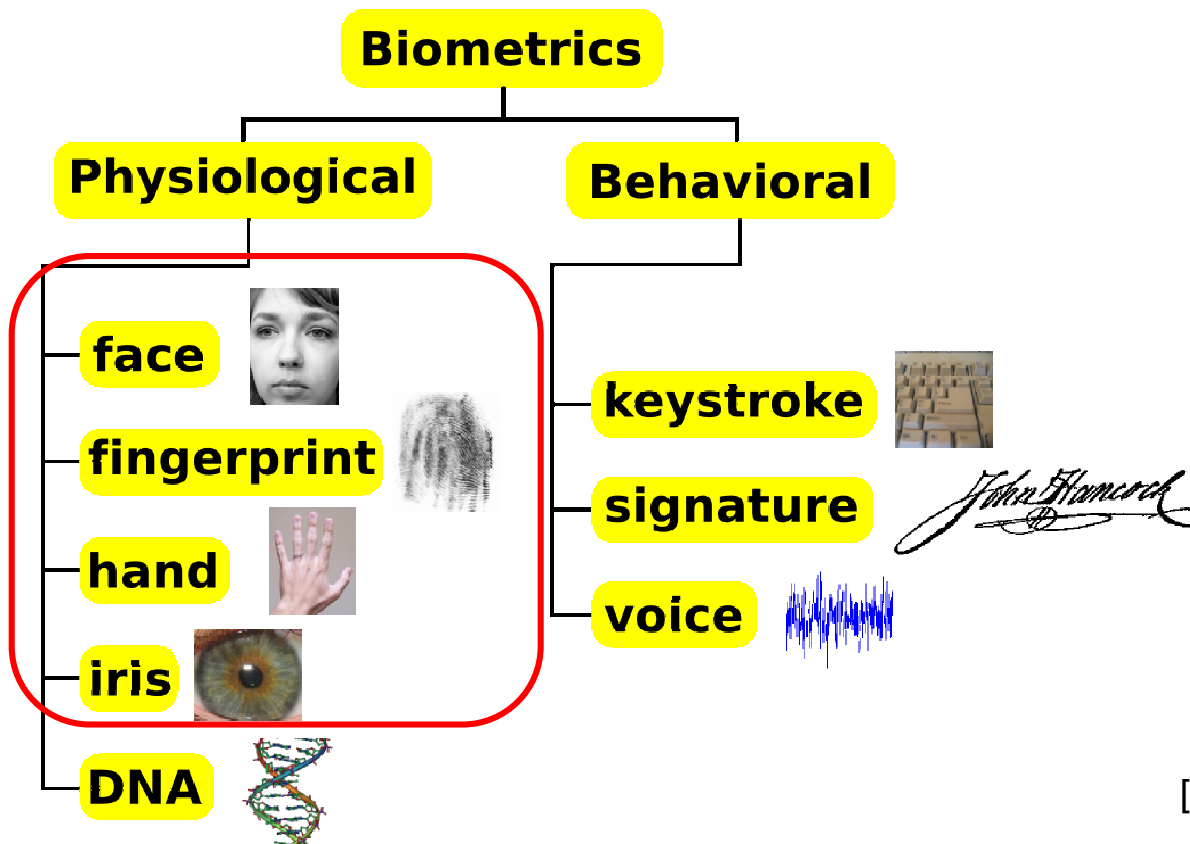
Head of Mathematics and Data Exploration Section (2008-9)

Science and Engineering Directorate

Canada

What is Biometrics?

Biometrics is an automated technique of measuring a physical characteristic (**biometric data**) of a person for the purpose of recognizing* him/her.



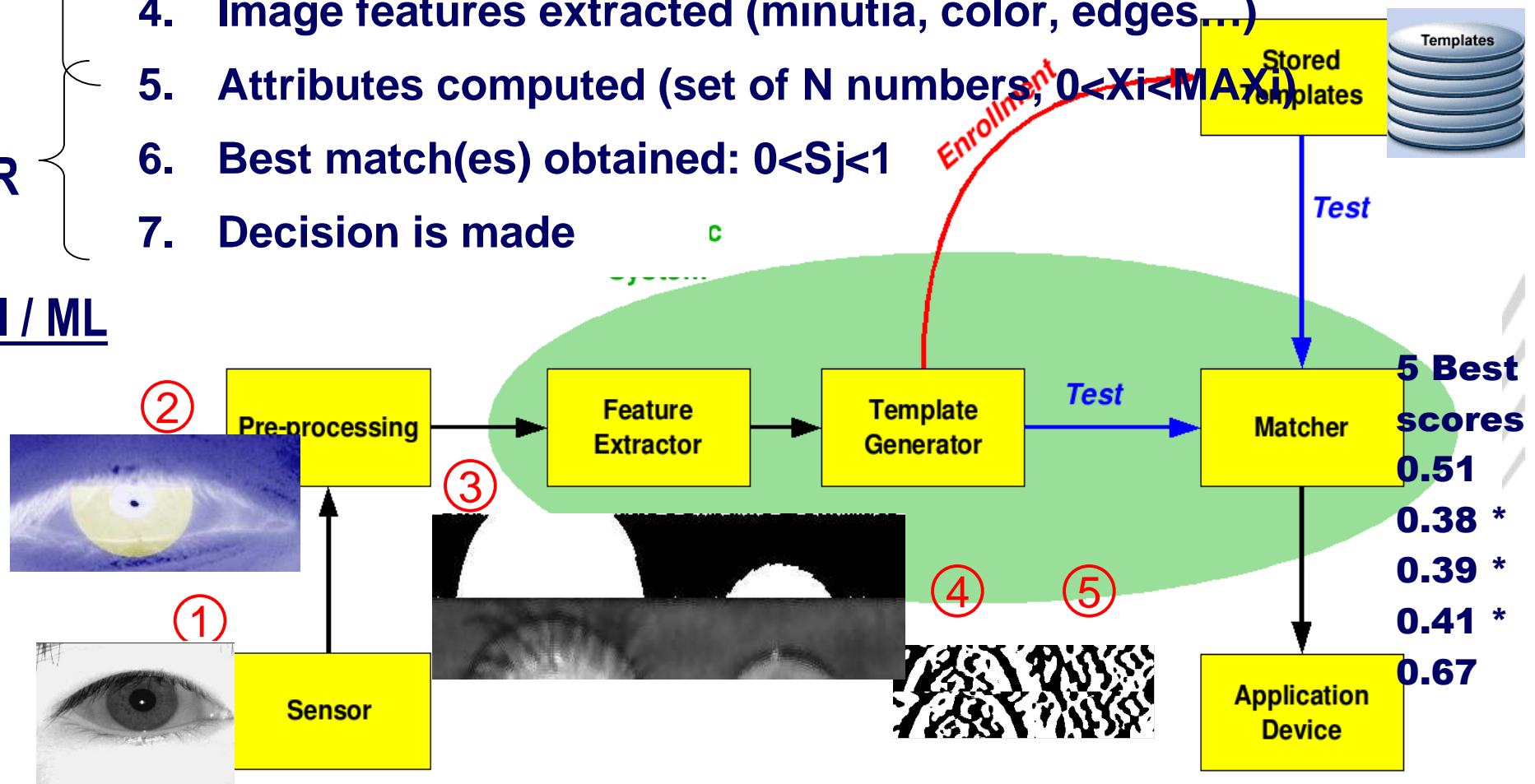
[CBSA NEXUS Iris Recognition system]

NB: Excellent (yet under-studied?) application of AI !

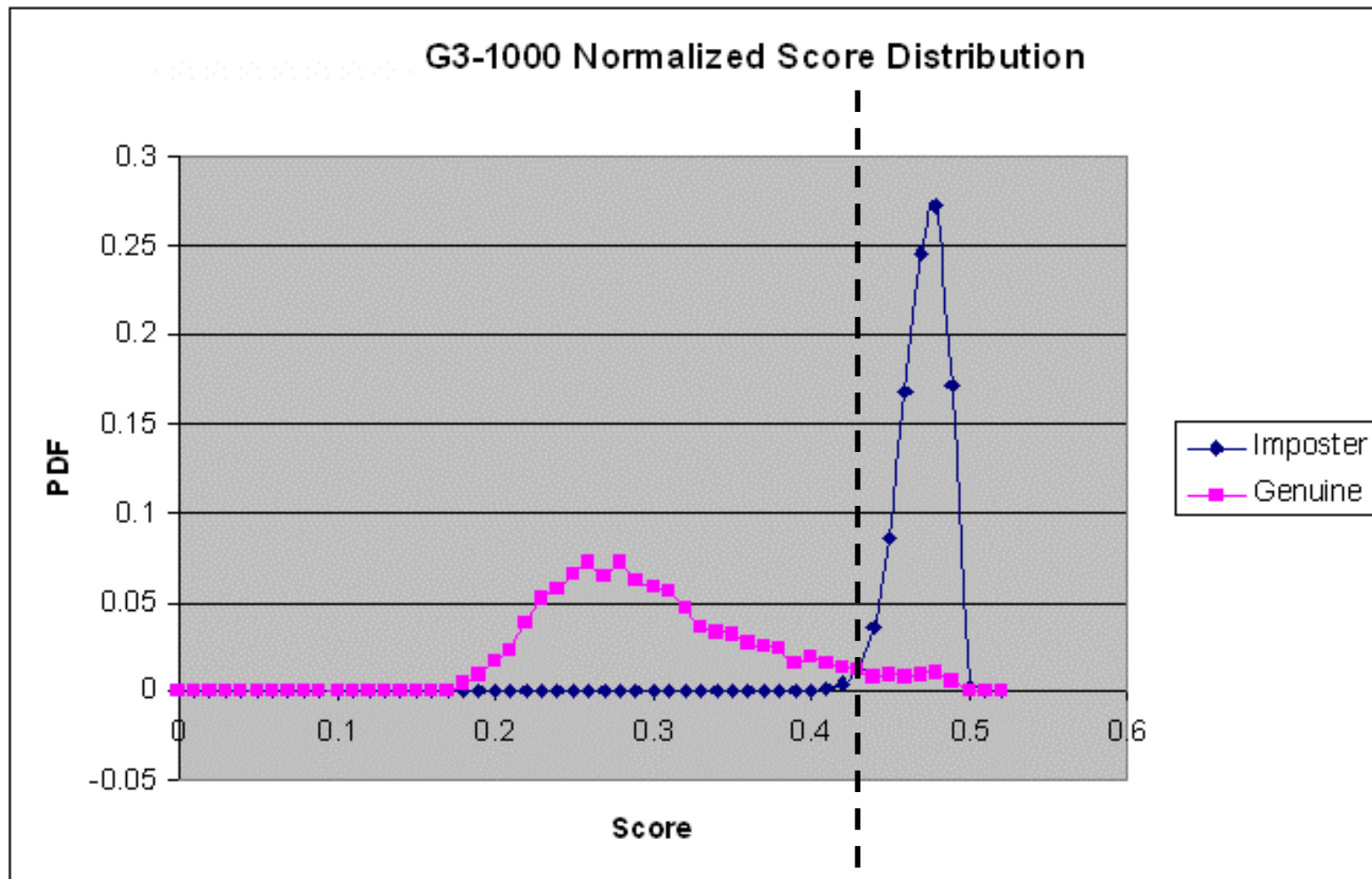
Biometric recognition process

- IP {
1. Image(s) captured
 2. Best image(s) selected and enhanced - preprocessing
 3. Biometric region extracted - segmentation
 4. Image features extracted (minutia, color, edges...)
- PR {
5. Attributes computed (set of N numbers, $0 < X_i < MAX_i$)
 6. Best match(es) obtained: $0 < S_j < 1$
 7. Decision is made

AI / ML



Scores for Genuine and Imposter users:



- NB: Traditionally, Match is when score is lower a threshold (ie. And it could be not the smallest score!)

CBSA - a prime user of Iris biometrics

Why iris ? – Easily accepted by public, touch-less / non-intrusive

Today: for collaborative user-engaged identification of pre-approved travellers in structured/overt environment (NEXUS)

Tomorrow: for fully-automated stand-off (on-the-fly) identification of Good and Bad people as they cross the border ? (3 persons crossing / sec)

Recent RFI examination (Feb 2009-Aug 2009) exposed the problems even with Today's systems/data

With Tomorrow's stand-off systems, these problems will be even more significant!

Gorodnichy, D. O. *“Evolution and evaluation of biometric systems”* IEEE Symposium: Computational Intelligence for Security and Defence Applications, Ottawa June 2009

Gorodnichy, D. O. *“Multi-order analysis framework for comprehensive biometric performance evaluation”*, SPIE Conf. on Defense, Security, and Sensing. Orlando, April 2010

Adding more AI to Biometrics...



ISO STANDARD 19795-1 (Biometric performance testing and reporting — Part 1: Principles and framework)

- Defines : Verification* as 1 to 1 Matching (below/above Threshold test) i.e. no 1 to N comparisons are done!
- Measures: Performance as False Match vs. False Non-Match Rates

Our goal: to bring AI back to the problem – to improve the performance!

- We call traditional (1 to 1) systems – Order 1 systems
- We do 1 to N matches to obtain the smallest score - Order 2 systems
- We recalibrate scores using the intrinsic properties of score distributions – Order 3 systems

* **Verification [ISO]:** application in which the user makes a positive claim to an identity, features derived from the submitted sample biometric measure are compared to the enrolled template for the claimed identity, and an accept or reject decision regarding the identity claim is returned.

Multi-order score analysis [2009]



Order 1 (Traditional):

- Single-score statistics (FMR/FNMR) and trade-off curves

Order 2:

- Examine all scores and report the best (smallest) score

Order 3:

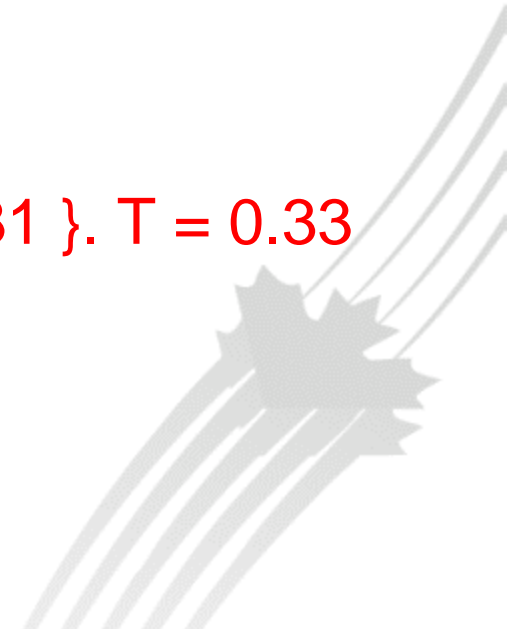
- Examine difference scores relationship

Five-score example: { 0.51, 0.32, 0.47, 0.34, 0.31 }. $T = 0.33$

Order 1 \rightarrow 0.32

Order 2 \rightarrow 0.31

But in reality it could have been 0.34 !



Goal: assign confidences to decisions

Given: Person X arrives at the kiosk and produces n scores:
n-tuple $S = (s_1, s_2, \dots, s_n)$, $s_i = \text{HD}(X, x_i)$

Find: Sequence of calibrated confidence scores:
the probability vector $C = (c_1, c_2, \dots, c_n)$, $c_i = P(\{X = x_i\} | S)$

How: as in probabilistic weather forecasting [DeGroot1983]

1. Make use of (assume) binomial nature of Genuine and Imposter score distributions [Daugman1993,2004]:

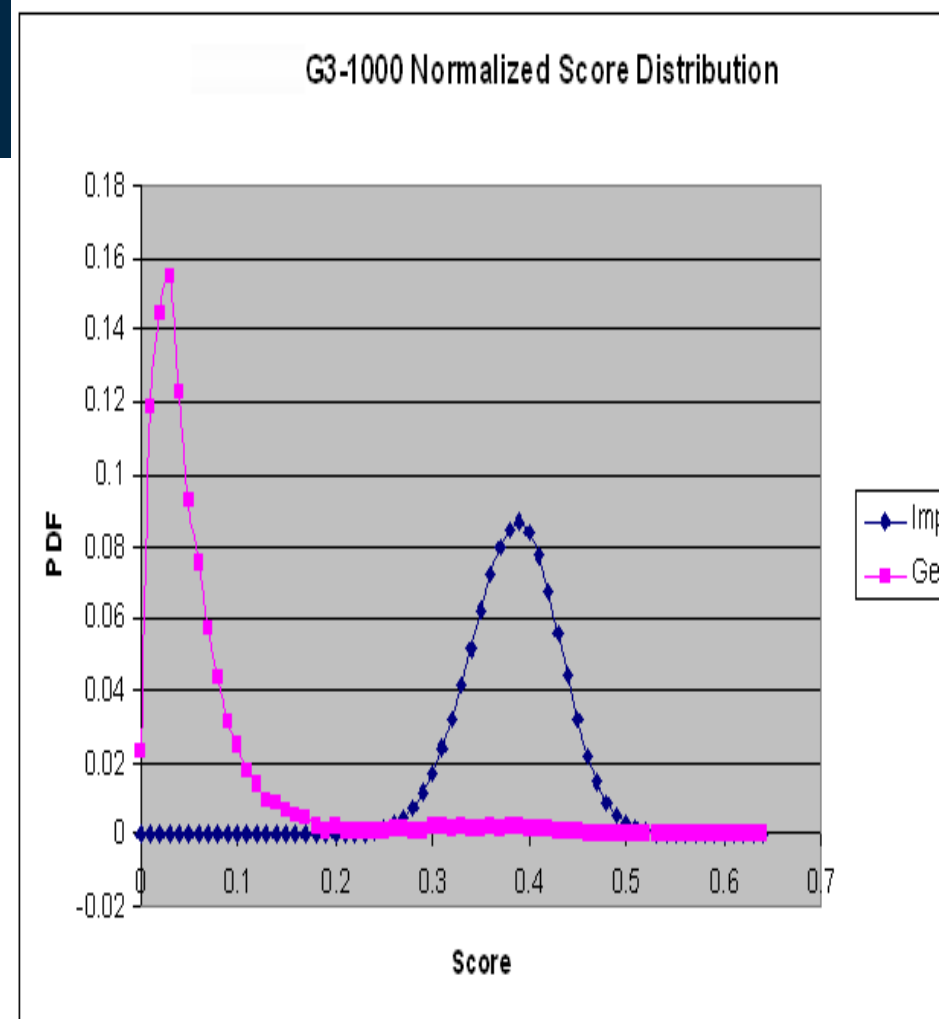
- $G \sim \text{Binom}(m', u')$, with $u' = 0.11$, $d' = 0.065$ ($m' \approx 115$).
- $I \sim \text{Binom}(m, u)$, with $u = 0.5$, $m = 249$ ($d \approx 0.03$)
- $P(\text{HD}=k/m) = \binom{m}{k} u^k (1-u)^{m-k}$

2. Bayes's Theorem for $c_i = P(\{X = x_i\} | S) =$
 $= P(\{X = x_i\} \wedge S) = P(\{X = x_i\} \wedge S) / P(S) = \dots$

3. $P(\{X = x_i\} \wedge S) = \dots$

Iris biometrics

- Image converted to 2048 binary digits {0,1}
 - only small subsets of bits are mutually independent [1].
- Impostor HD scores follow binomial distribution:
 $I \sim \text{Binom}(m, u)$,
 $m = 249$ and $u = 0.5$.
- The variable m represents the degrees-of-freedom and is a function of the mean u and the standard deviation d :
 $m = u(1 - u) / d^2$



- Genuine HD scores [2]:
 $G \sim \text{Binom}(m', u')$ with
 $u' = 0.11$, $d' = 0.065$

Main theorem and proof:



Theorem 3.1 Let G be the set of genuine matching scores, and I be the set of impostor matching scores. Suppose $G \sim \text{Binom}(\hat{m}, \hat{u})$ and $I \sim \text{Binom}(m, u)$. Let $p_i = P(X = x_i)$ and $q = 1 - \sum_{i=1}^n p_i$. Let $S = (s_1, s_2, \dots, s_n)$ be the n -tuple of matching scores produced by person X . Then for each $1 \leq i \leq n$, we have

$$c_i = P(X = x_i | S) = \frac{p_i z_i}{\sum_{i=1}^n p_i z_i + q \cdot \frac{(1-u)^m}{(1-\hat{u})^{\hat{m}}}}, \quad \text{where } z_i = \frac{\binom{\hat{m}}{\hat{m}s_i}}{\binom{m}{ms_i}} \cdot \left(\frac{\hat{u}^{\hat{m}}(1-u)^m}{u^m(1-\hat{u})^{\hat{m}}} \right)^{s_i}.$$

Proof: For each $1 \leq i \leq n$, define $r_i = P(\{X = x_i\} \wedge S)$. Also define $r_{imp} = P(\{X \notin \{x_1, x_2, \dots, x_n\}\} \wedge S)$.

By definition, $r_{imp} = P(S) - \sum_{i=1}^n r_i$. By Bayes' Theorem, we have

$$c_i = P(\{X = x_i\} | S) = \frac{P(\{X = x_i\} \wedge S)}{P(S)} = \frac{r_i}{r_1 + r_2 + \dots + r_n + r_{imp}}.$$

To calculate $r_i = P(\{X = x_i\} \wedge S)$, we multiply the probabilities of the following $n + 1$ independent events: it is x_i who comes to the kiosk; the genuine matching score $HD(X, x_i)$ is s_i ; and the impostor matching score $HD(X, x_j)$ is s_j for all $1 \leq j \leq n$ with $j \neq i$.

Since $G \sim \text{Binom}(\hat{m}, \hat{u})$, there are \hat{m} degrees-of-freedom, and the probability that any of these \hat{m} bits differ is \hat{u} . So if $HD(X, x_i) = s_i$, then $\hat{m}s_i$ of the \hat{m} bits differ. We derive the analogous result for the impostor distribution $I \sim \text{Binom}(m, u)$, for all $1 \leq j \leq n$ with $j \neq i$. Therefore, we have

$$r_i = p_i \binom{\hat{m}}{\hat{m}s_i} \hat{u}^{\hat{m}s_i} (1-\hat{u})^{\hat{m}-\hat{m}s_i} \cdot \prod_{j=1, j \neq i}^n \binom{m}{ms_j} u^{ms_j} (1-u)^{m-ms_j}$$

Simple example to illustrate



Enrolled: three individuals $\{x_1, x_2, x_3\}$, six bits in iris string.

- Thus, $n = 3$, $m = m' = 6$.
- $G = \text{Binom}(m', u')$, $I = \text{Binom}(m, u)$ with $u' = 1/3$ and $u = 1/2$.
- $x_1 = [0, 1, 0, 1, 0, 1]$, $x_2 = [1, 0, 0, 1, 1, 1]$, $x_3 = [1, 0, 1, 1, 0, 1]$

New person: $X = [0, 1, 0, 1, 0, 1]$.

- Matching scores $S = (0, 0.5, 0.5)$. Decision scores: $(1, 0, 0)$.

Using the theorem (for $q=0$ and $P_1=P_2=P_3$), we obtain:

- confidence scores $C = (0.8, 0.1, 0.1)$.

How to apply to real system?

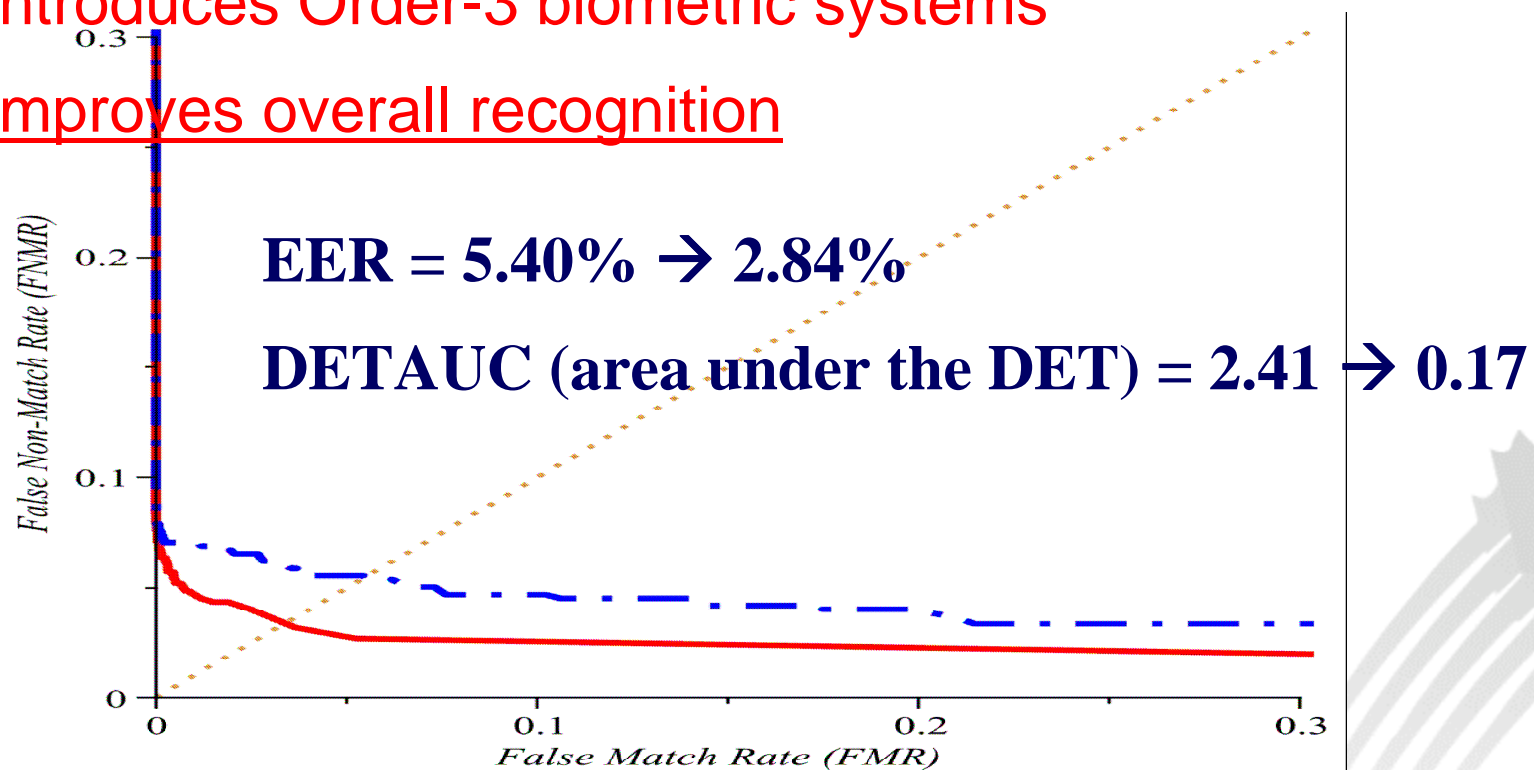
- Vendor should provide: m', u' m, u
- User knows: P_i, q (a-priory probabilities of each person / imposter)
- Applied to best 5 scores.



Applied to real system

Proposed probabilistic score calibration can be added to any system at little computation cost as post-processing filter:

- Provides more meaningful output - for risk mitigating procedures
- Introduces Order-3 biometric systems
- Improves overall recognition



Theoretical Proof [AI 2010]



Theorem: If G and I are both binomially distributed, then the algorithm whose scores and match decisions are based on the calibrated confidence function (Eq. 1), rather than on the matching scores, produces the biometric system's best possible DET curve both at the score level and at the decision level.

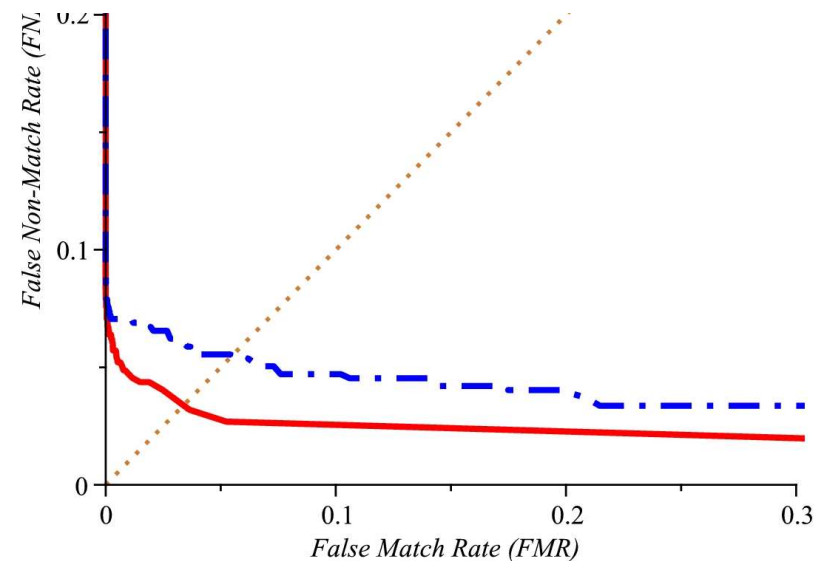
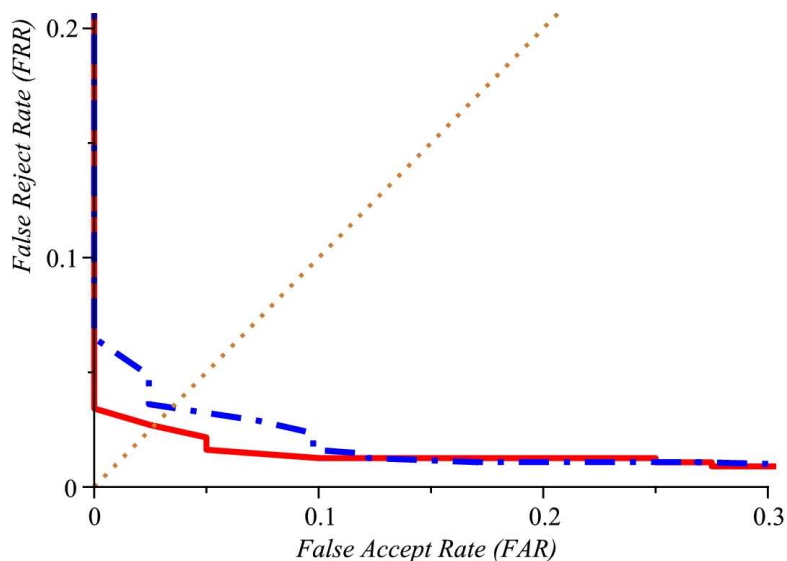


(1 to 1) vs. (1 to N) systems



1 to 1 (Order 1) systems: DET curve at the score level graphs the false match rate (FMR) against the false nonmatch rate (FNMR) over all possible thresholds, which is done by examining the scores given to genuine and impostor comparisons. On the other hand, the

1 to N (Order 2) systems: DET curve at the decision level graphs the false accept rate (FAR) against the false reject rate (FRR) over all possible thresholds, which is done by comparing all n scores and seeing if the highest score lies above the threshold.



Invitation to further this work



- Apply to systems with non-binomial score distributions (eg Normal)
 - Derive theoretical formulas and approximations
- Contribute to ISO/industry definitions, standards, practices – to make better use of AI in Biometrics
 - through the application of multi-order score analysis
- Travel often? – Use NEXUS !
- Your BIG biometric R&D partner and user is right here in Ottawa
 - Project with École de Technologie Supérieure (Montreal) on biometric decision fusion for finding asymptotic biometric performance bounds
- Join ISO SC-37 (Biometrics) group!
 - It's free and they need AI experts

Contact: [Dmitry Gorodnichy \(dmitry.gorodnichy@cbsa.gc.ca\)](mailto:dmitry.gorodnichy@cbsa.gc.ca)