

Relating ROC and CMC Curves

Work done with Brian DeCann

Arun Ross

Associate Professor
Michigan State University

<http://www.cse.msu.edu/~rossarun>

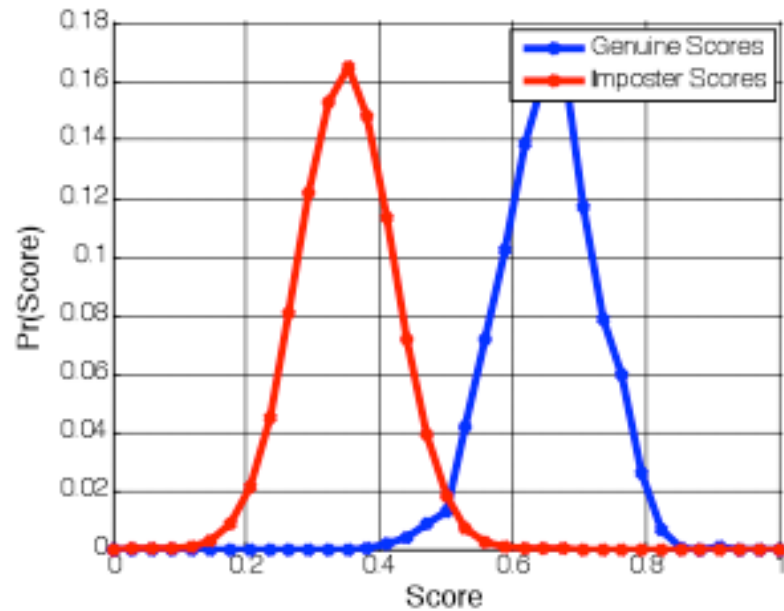
Introduction

- Performance of a **verification** system is summarized using Receiver Operating Characteristic (**ROC**) curve
- Performance of a **closed-set identification** system is summarized using Cumulative Match Characteristic (**CMC**) curve
- Can the CMC curve be **derived** from the ROC curve and vice-versa?

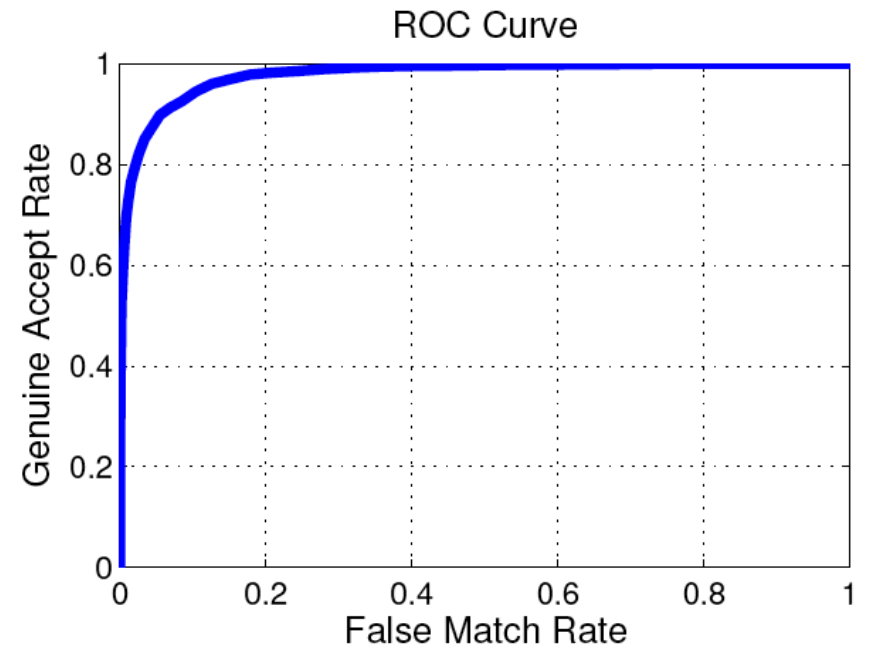
ROC Curve

- Biometrics samples are **compared** against each other
- **Genuine** and **impostor** scores are generated
- False Match Rate (**FMR**) and False Non-match Rate (**FNMR**) are computed at multiple **thresholds**
- ROC Curve: True Match Rate versus False Match Rate
- **ROC Curve: Aggregate Statistics**

ROC Curve



Match Score Distributions

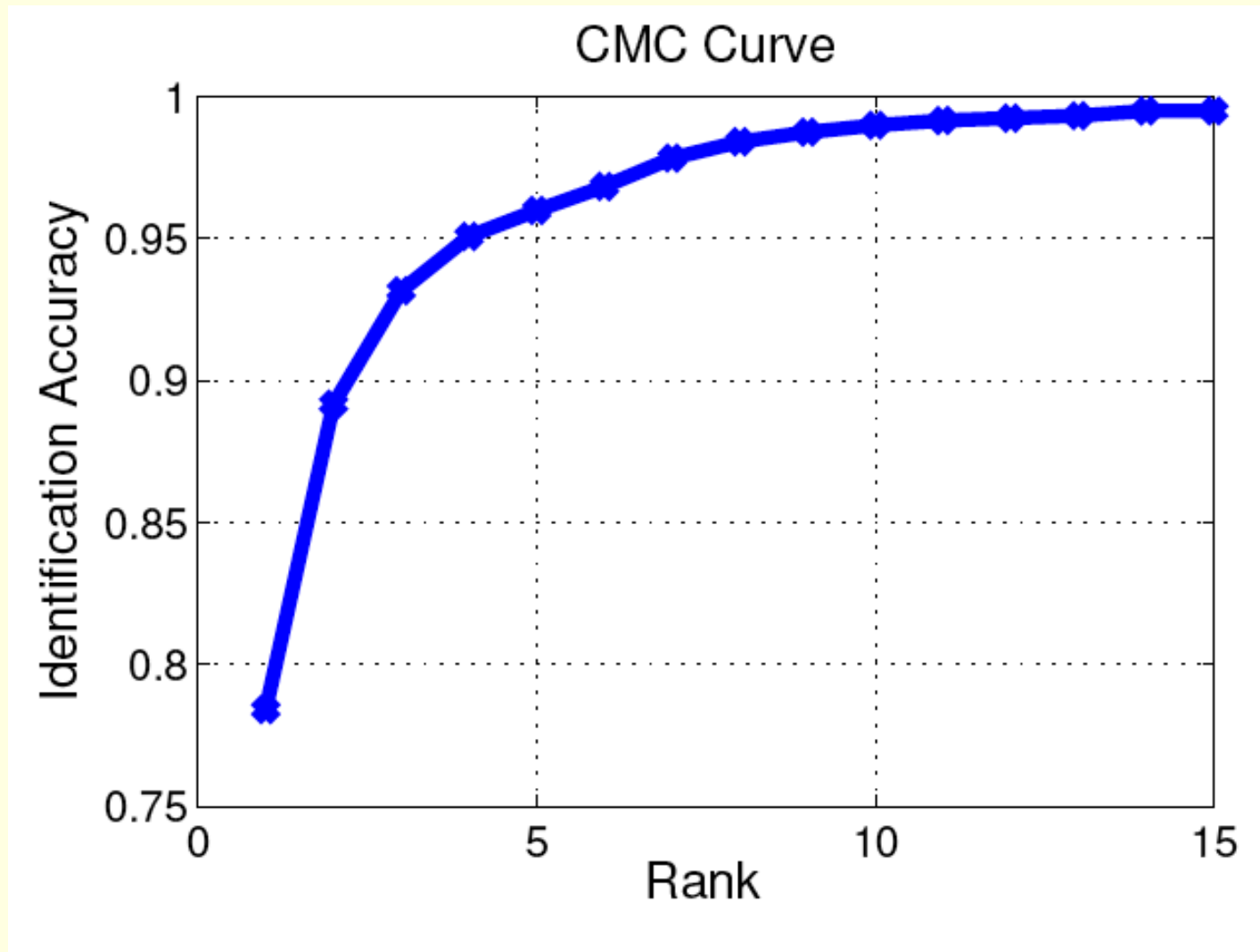


ROC Curve

CMC Curve

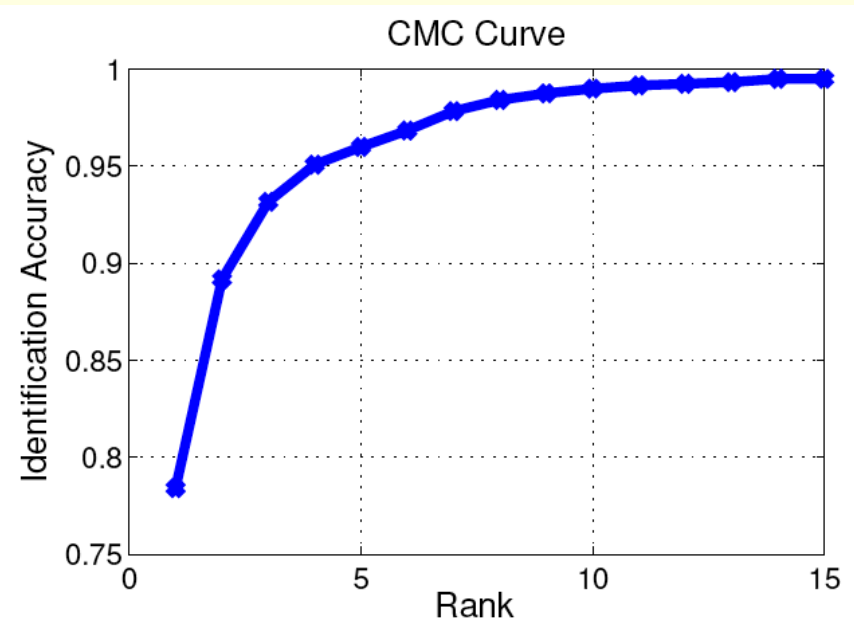
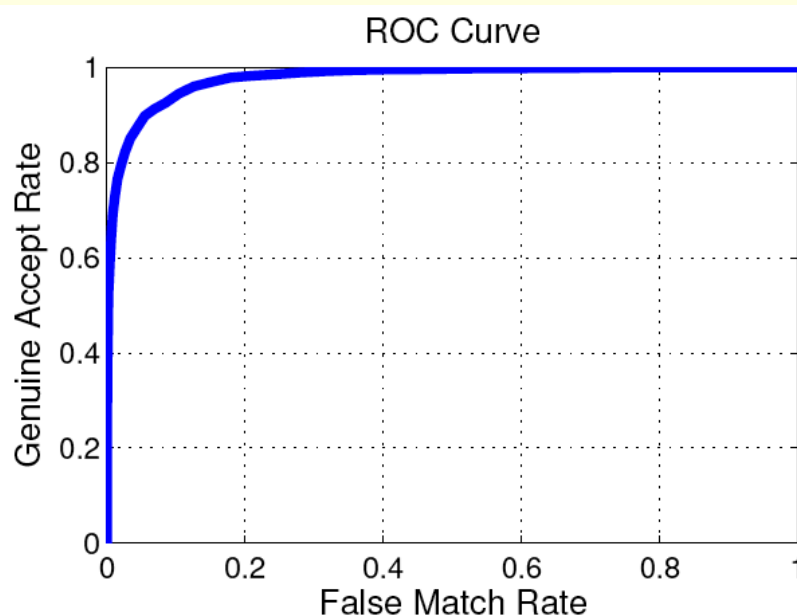
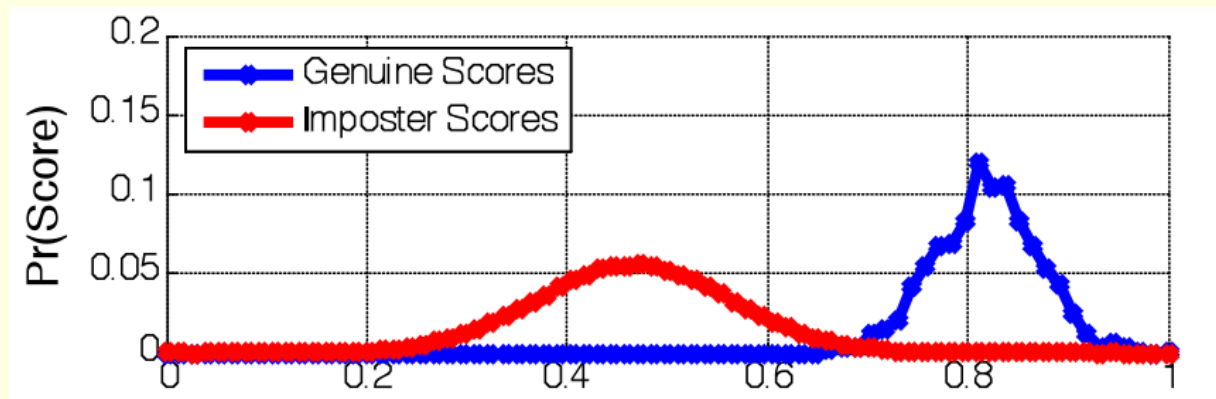
- Each probe biometric sample is **compared** against all gallery samples
- The resulting scores are **sorted** and **ranked**
- Determine the rank at which a **true match** occurs
- True Positive Identification Rate (**TPIR**): Probability of observing the **correct identity** within the **top K ranks**
- CMC Curve: Plots TPIR against ranks
- **CMC Curve: Rank-based metric**

CMC Curve



CMC versus ROC

- It is reasonable to expect a **good ROC curve** to be associated with a **good CMC curve** and vice-versa



Predicting CMC from ROC

- The CMC can be **predicted** from the ROC data
 - Bolle et. al. (2005), Hube (2006)

Bolle

$$Rank - n = \sum_{k=1}^n \binom{N-1}{k-1} \int_0^{\infty} F_G(s) FAR(s)^{k-1} (1 - FAR(s))^{(N-k)} ds$$

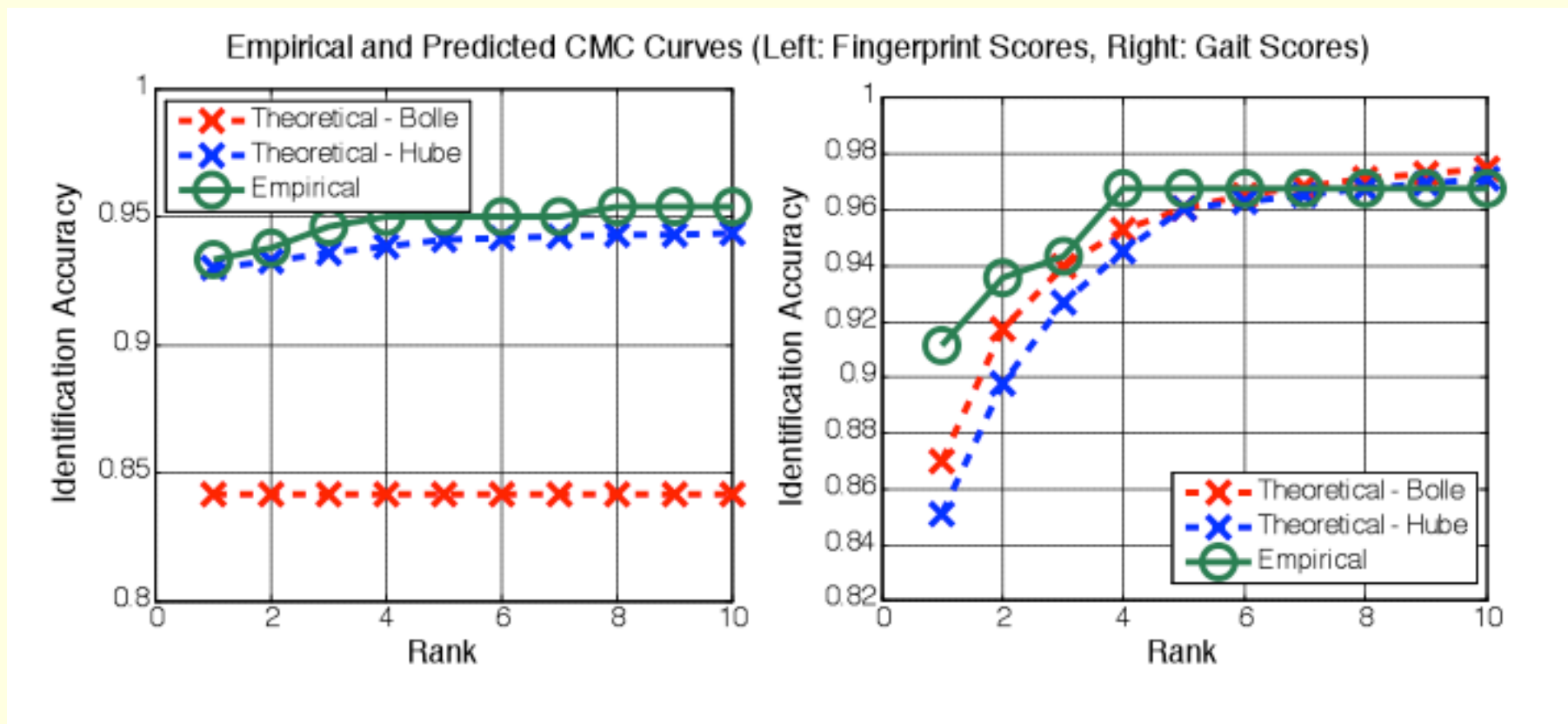
Hube

$$Rank - n = TPR \left(FAR = \frac{n}{N} \right), \text{ where } TPR = (1 - FNMR)$$

- R. Bolle, J. Connell, S. Pankanti, N. Ratha, and A. Senior. *The Relation Between the ROC Curve and the CMC*. AutoID 2005
- J. Hube. *Using Biometric Verification to Estimate Identification Performance*. BSYM 2005

Predicting CMC from ROC

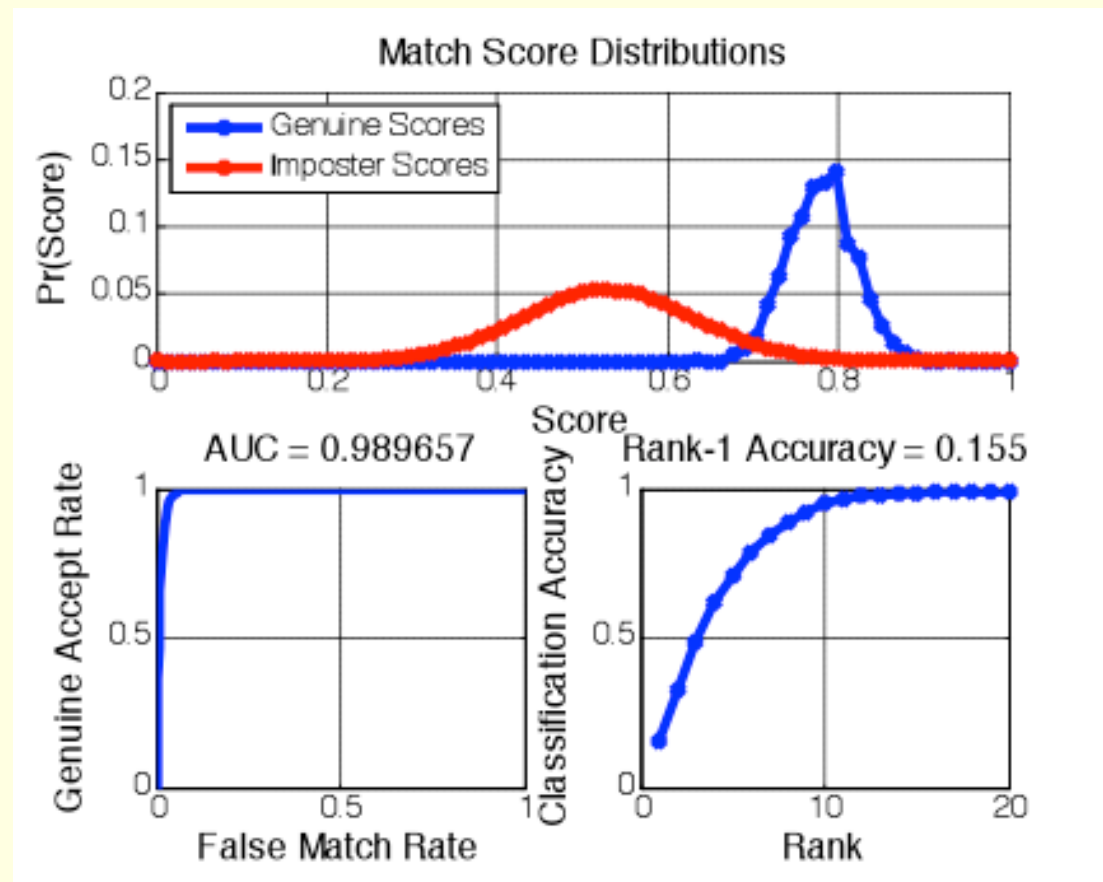
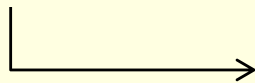
- But neither model **perfectly** predicts the empirical CMC curve



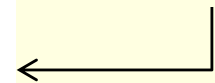
ROC versus CMC

- DeCann and Ross (2012) showed that it is possible for a **good ROC** curve to be associated with a **poor CMC** curve and vice-versa

**GOOD ROC
CURVE**

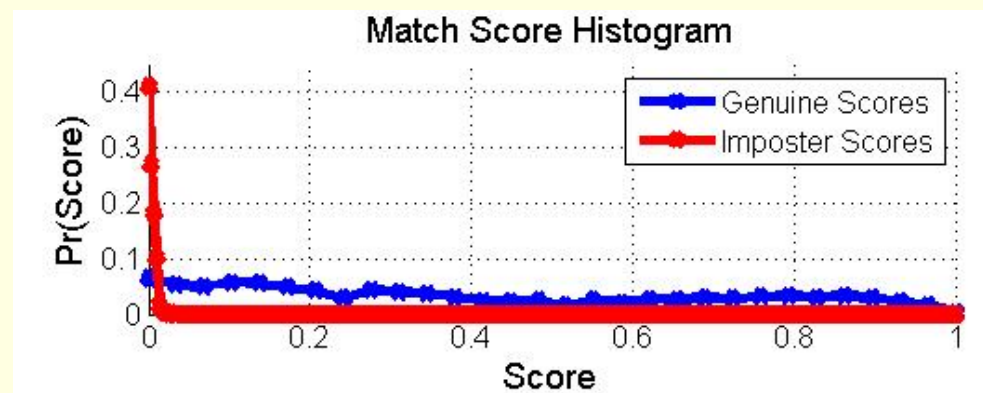
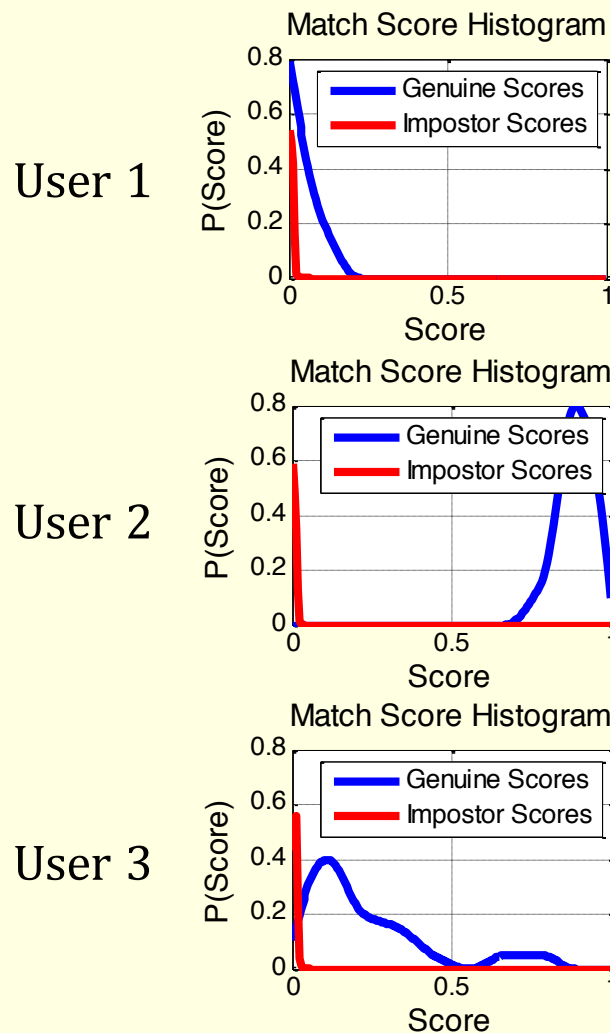


**POOR CMC
CURVE**

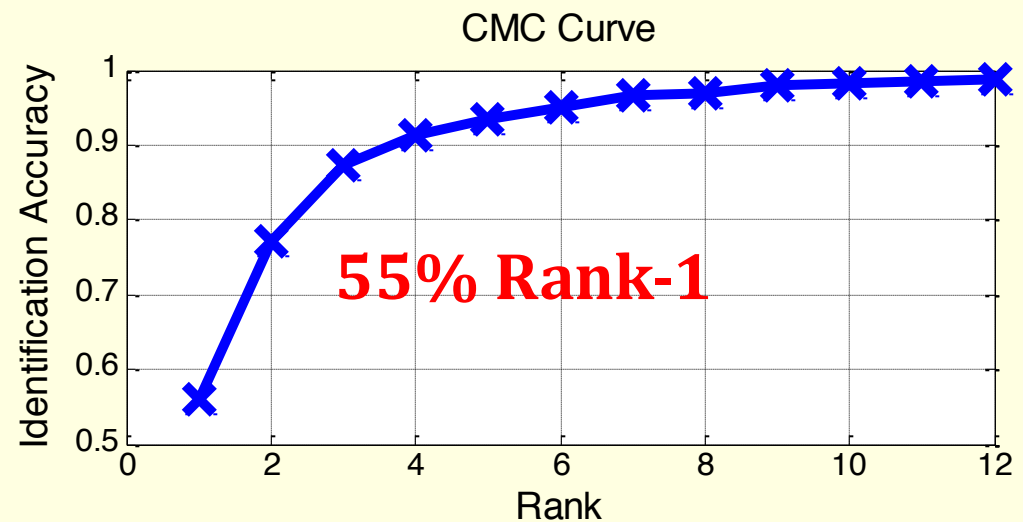
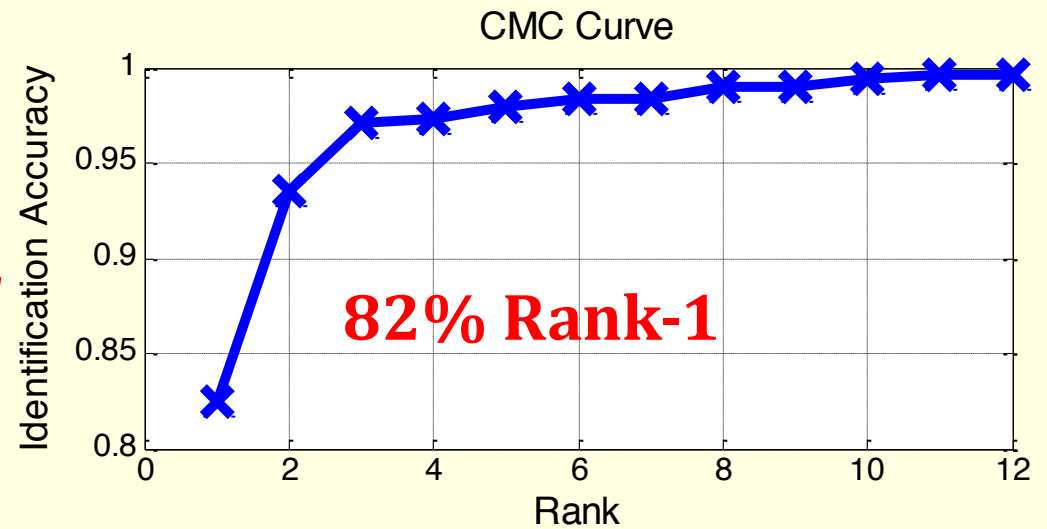
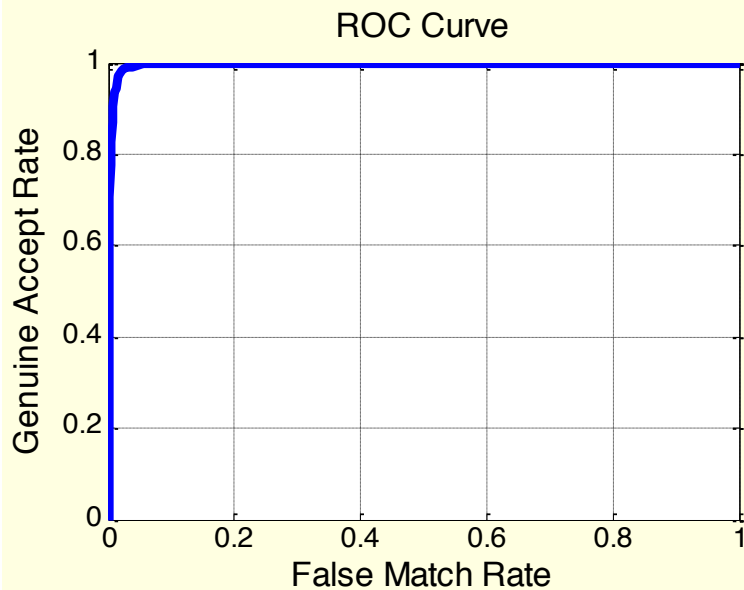


Why did CMC prediction models fail?

- Each identity contributes uniquely to the system performance, i.e., **aggregate** statistics do not necessarily predict **ranked statistics**



One ROC Curve: Multiple CMC Curves

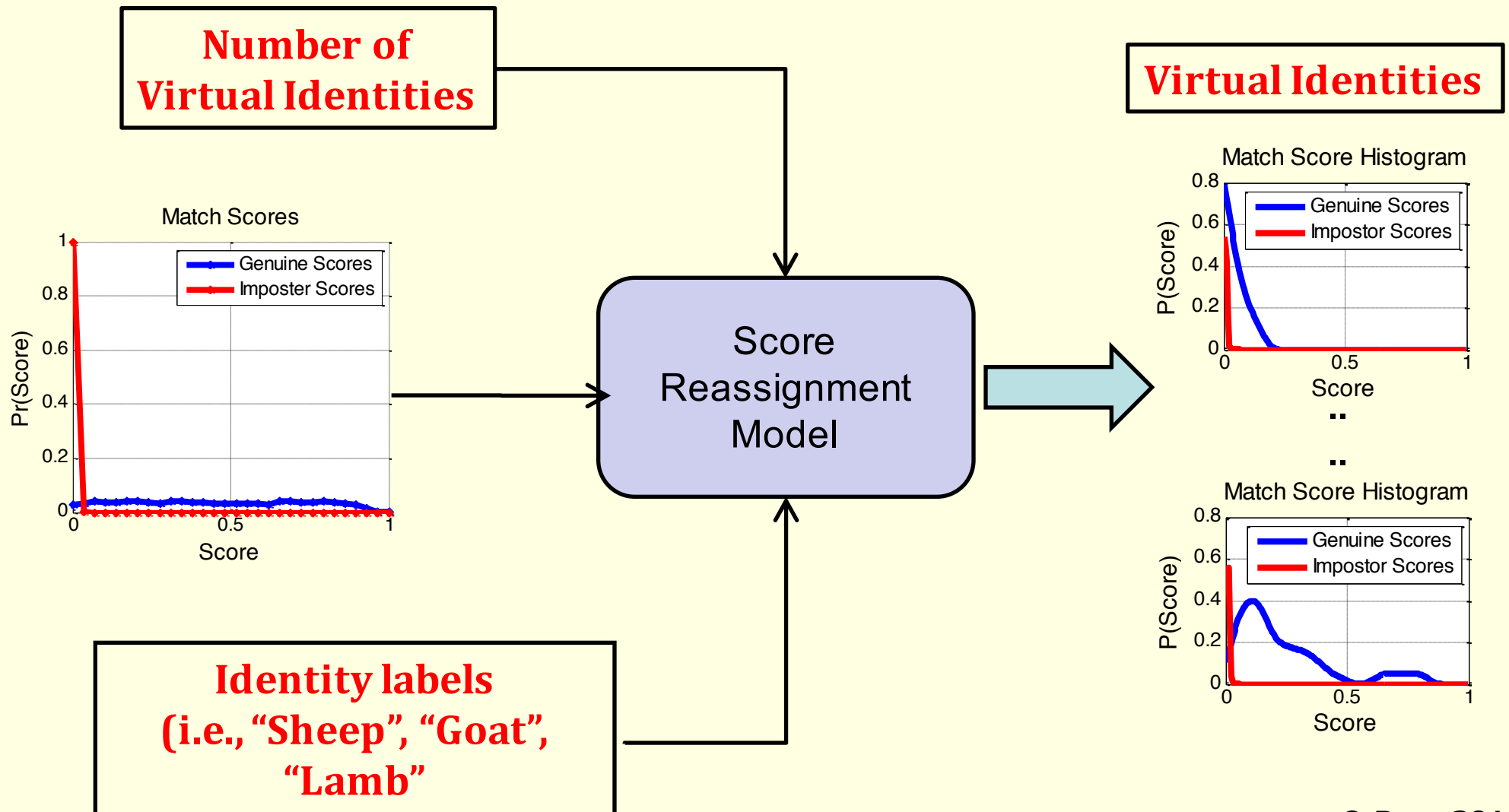


Virtual Identities

- **Input:** Set of genuine and impostor match scores
- **Output:** Virtual identities with different rank-based statistics
- **Method:** “Reassign” match scores to virtual identities according to the “Dodgington’s Zoo” concept
 - **Sheep:** Low FMR and FNMR
 - **Goats:** High FNMR
 - **Lambs:** High FMR

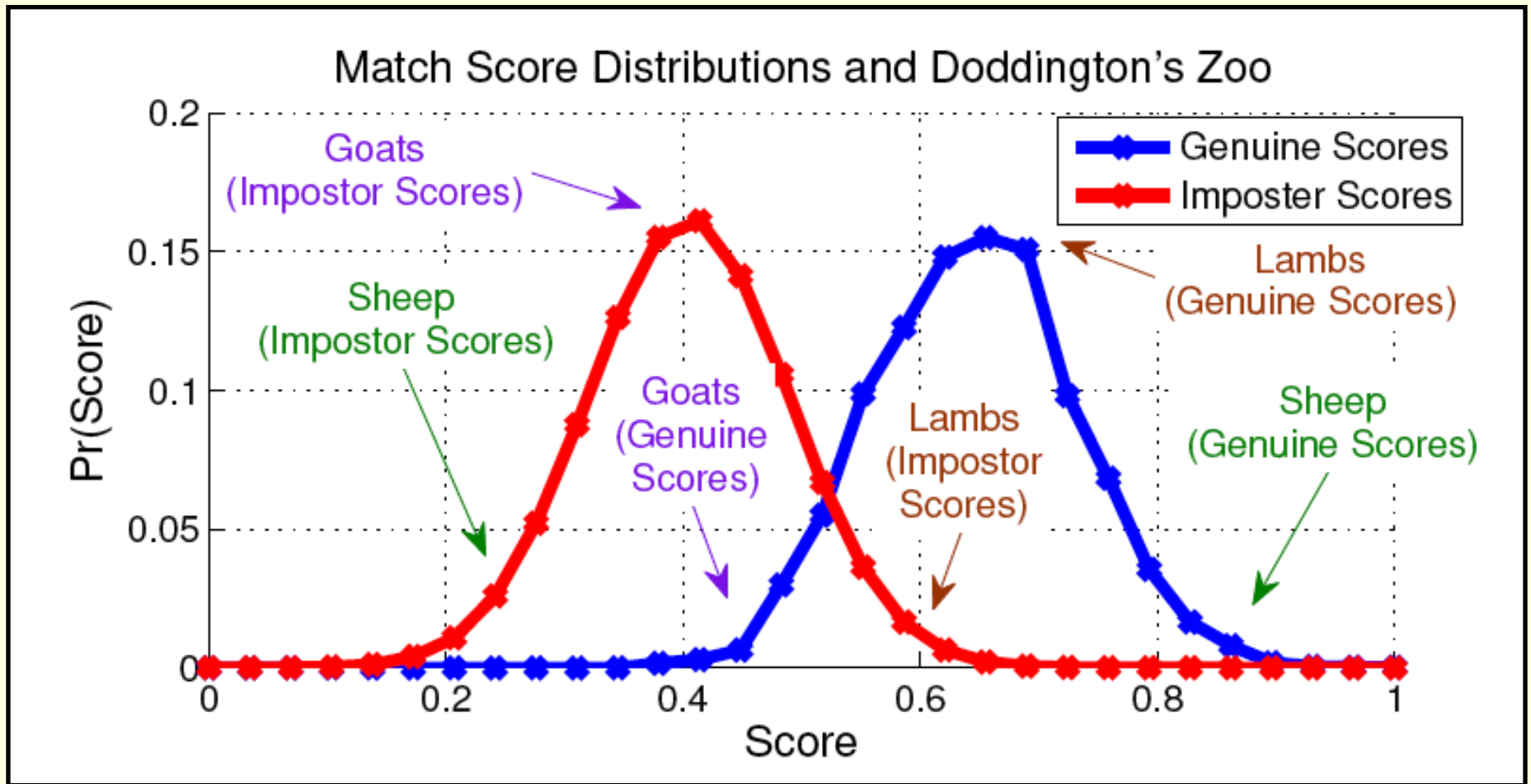
Reassigning Match Scores

- Set of genuine and impostor match scores



Sampling Match Scores

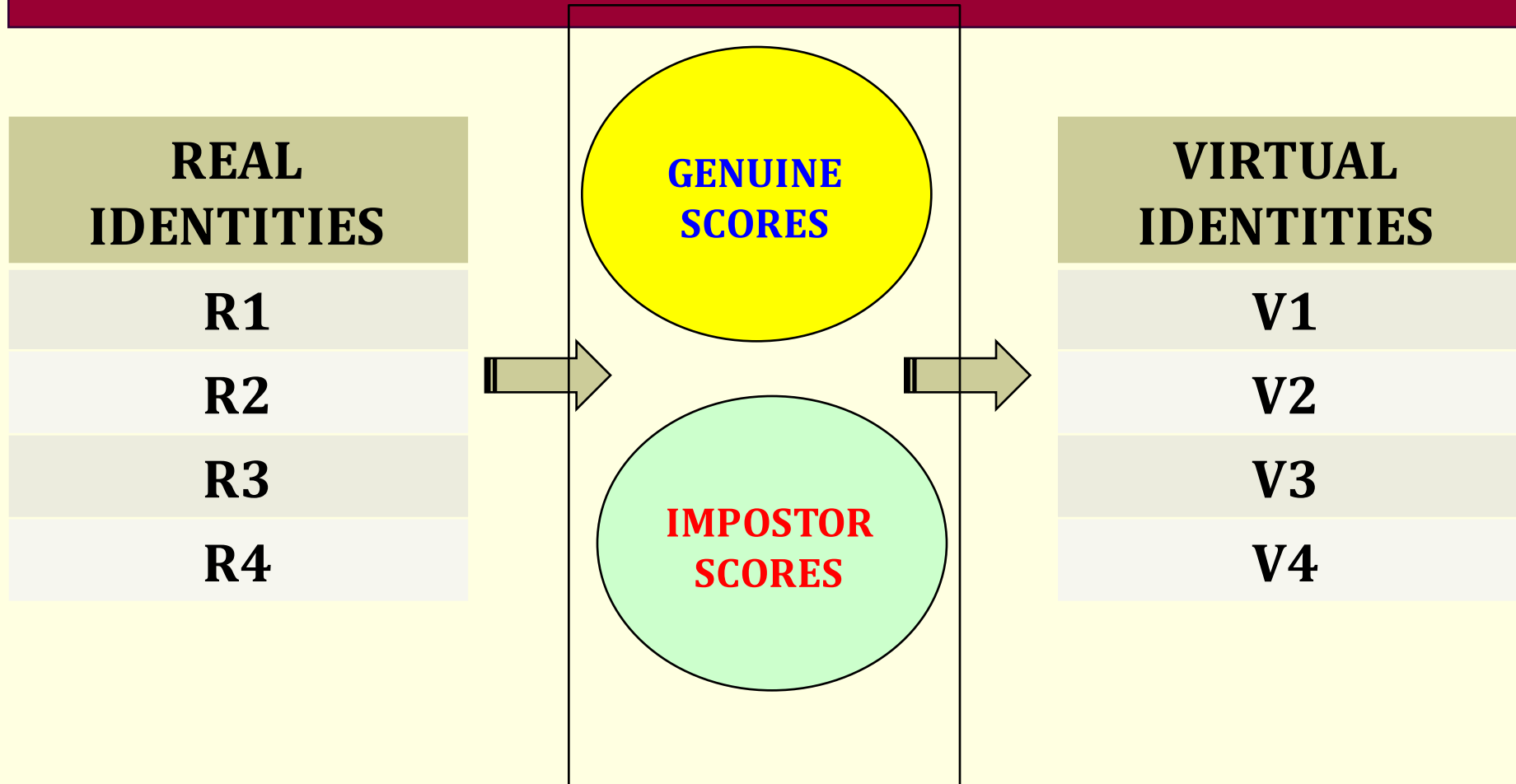
- Depending upon “Sheep”, “Goat”, “Lamb” labels



Sampling Rationale

- **Genuine Scores:** Use the label (“Sheep”, “Goat”, “Lamb”) to assign genuine match scores to a virtual identity
- **Impostor Scores:** Use the labels of “**pairs**” of virtual identities to assign impostor match scores to a virtual identity

From Real to Virtual



- Aggregate Statistics do **not** change

Reassigning Genuine Scores

Algorithm 1: Reassigning Genuine Scores

Input: Vector \mathbf{s}_{Gen} , containing the genuine scores.

Vector χ , a set containing the labels of each identity
(e.g., “Sheep”, “Goat”, “Lamb”).

Define: δ, ϵ_{Gen} : Scaling parameters.

Output: Matrix \mathbf{S} populated with genuine scores.

\\ begin algorithm

Step 1: For each identity, note the assigned label.

Step 2a: Draw a genuine score (without replacement), ϕ ,
 \mathbf{s}_{Gen} , from within subset \mathbf{s}_{rng} , where

$$\mathbf{s}_{rng} = (\mu_{Gen} + \sigma_{Gen}, 1), \text{ if } \chi_n = \textit{Sheep}.$$

$$\mathbf{s}_{rng} = (0, \mu_{Gen} - \sigma_{Gen}), \text{ if } \chi_n = \textit{Goat}.$$

$$\mathbf{s}_{rng} = (0, \mu_{Gen} + \sigma_{Gen}), \text{ if } \chi_n = \textit{Lamb}.$$

Reassigning Genuine Scores

Step 2b: If \mathbf{s}_{rng} is a null set, and $\mathbf{s}_{rng} = (a, b)$,
set $a = \delta \cdot a$, $b = \frac{b}{\delta}$ and repeat Step 2a.

Step 3a: Draw $\binom{N_G}{2} - 1$ scores (without replacement)
from \mathbf{s}_{Gen} within $\phi \pm \epsilon_{Gen}$.

Step 3b: If less than $\binom{N_G}{2} - 1$ scores can be drawn
set $\epsilon_{Gen} = \frac{\epsilon_{Gen}}{\delta}$ and repeat Step 3a.

Step 4: Store the sampled genuine scores in \mathbf{S} .

return \mathbf{S}

\\ end algorithm

Reassigning Impostor Scores

Algorithm 2: Reassigning Impostor Scores

Input: Vector \mathbf{s}_{Imp} , containing the impostor scores.

Matrix \mathbf{S} , where sampled genuine scores are stored (from Alg. 1) and sampled impostor scores will be stored.

Vector χ , containing the labels of each identity (e.g., “Sheep”, “Goat”, “Lamb”).

$\mathbf{S}_{Gen}^n, \mathbf{S}_{Gen}^m$, Assigned genuine scores for identities n, m .

Define: δ, ϵ_{Imp} : Scaling parameters.

Output: Matrix \mathbf{S} populated with genuine and impostor scores.

\\ begin algorithm

Step 1: For all combinations of n and m ($n = 1, \dots, N$,
 $m = n + 1, \dots, N$), note χ_n and χ_m .

Reassigning Impostor Scores

Step 2: Draw an impostor score, ϕ from \mathbf{s}_{Imp} , within interval \mathbf{s}_{rng} , where

$\mathbf{s}_{rng} = (0, \min\{\max\{\mathbf{S}_{Gen}^n\}, \max\{\mathbf{S}_{Gen}^m\}\})$,
if $\chi_n = \textit{Sheep}$ or \textit{Goat} , $\chi_m = \textit{Sheep}$ or \textit{Goat} .

$\mathbf{s}_{rng} = (0, \max\{\mathbf{S}_{Gen}^n\})$,
if $\chi_n = \textit{Sheep}$ or \textit{Goat} , $\chi_m = \textit{Lamb}$.

$\mathbf{s}_{rng} = (0, \max\{\mathbf{S}_{Gen}^m\})$,
if $\chi_n = \textit{Lamb}$, $\chi_m = \textit{Sheep}$ or \textit{Goat} .

$\mathbf{s}_{rng} = (0, 1)$, if $\chi_n = \chi_m = \textit{Lamb}$.

Step 3: If \mathbf{s}_{rng} is a null set, $\mathbf{s}_{rng} = (0, 1)$.

Reassigning Impostor Scores

Step 4a: Draw $N_G^2 - 1$ scores from \mathbf{s}_{Imp} within $\phi \pm \epsilon_{Imp}$.

Step 4b: If less than $N_G^2 - 1$ scores can be drawn

set $\epsilon_{Imp} = \frac{\epsilon_{Imp}}{\delta}$, and repeat Step 4a.

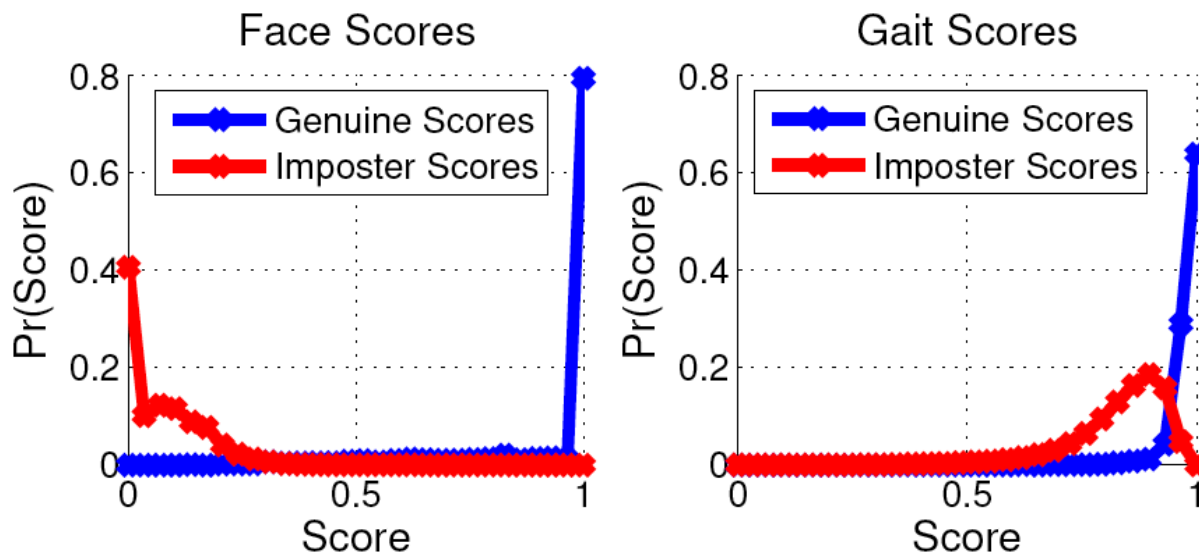
Step 5: Store the sampled impostor scores in \mathbf{S} .

return \mathbf{S}

\\ end algorithm

Datasets Used

- **Face:** WVU Multimodal Dataset
 - 240 subjects, 5 Samples / subject
 - Match scores computed using VeriLook
- **Gait:** CASIA B dataset
 - 124 subjects, 6 samples / subject
 - Match scores computed using Gait Curves algorithm



Evaluation Criteria

- ROC data: Area underneath the ROC (AUC)
- CMC data: Weighted Rank-M strategy

Generate Virtual Identities

- Generate virtual identities with different input parameters: (% Sheep, % Goats, % Lambs)
- Compute AUC and Rank-M values

Sheep (%)	Goat (%)	Lambs (%)	AUC (Face)	Rank-M (Face)	AUC (Gait)	Rank-M (Gait)
100	0	0	0.999	1.0	0.980	1.0
82	10	8	0.999	1.0	0.980	0.966
50	26	24	0.999	0.997	0.980	0.915
15	10	75	0.999	0.997	0.980	0.800

Same Aggregate Statistics

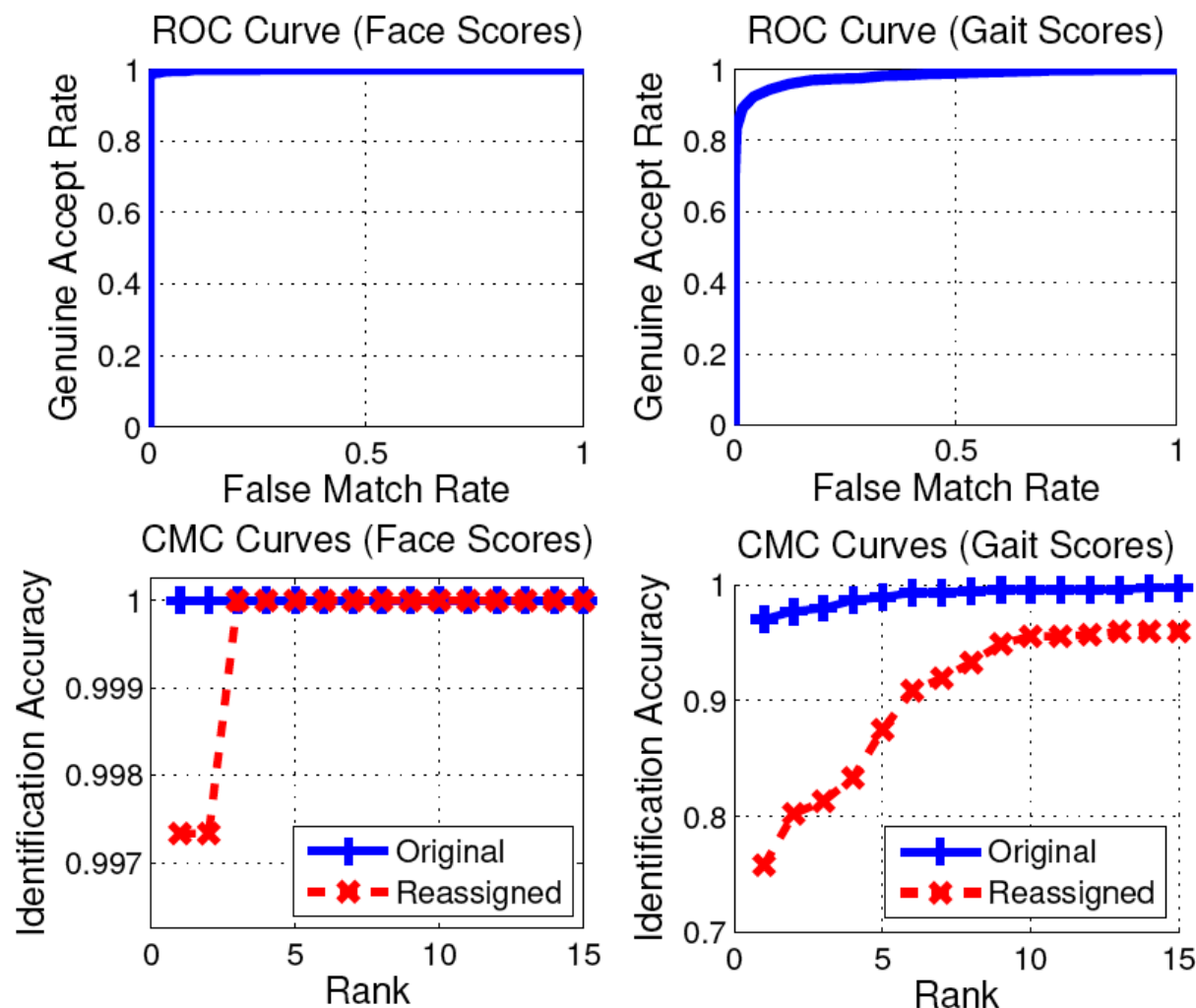
Different Rank Statistics

Note: Increasing the proportion of Goats or Lambs *decreases* Rank-M performance

A Closer Look

- ROC and CMC curves for “Original” and “Reassigned”
- (15% Sheep, 10% Goats, 75% Lambs)

100% to
99.7%



99% to
75%

Summary

- It is possible for a **single ROC** curve to be associated with **multiple CMC** curves
- The distribution of “**Sheep**”, “**Goat**”, “**Lamb**” in the target population results in this phenomenon
- Any **ROC-CMC prediction** model, should account for this variability in user performance
- **Soft biometric** traits are more likely to exhibit this type of disparity
- Reporting both **ROC** and **CMC** curves is recommended
- **Note: Closed-set identification**

Project sponsored by ONR

Reading Material

- B. DeCann and A. Ross, "**Relating ROC and CMC Curves via the Biometric Menagerie**," Proc. of 6th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), (Washington DC, USA), September 2013
- B. Decann and A. Ross, "**Can a Poor Verification System be a Good Identification System? A Preliminary Study**," Proc. of IEEE International Workshop on Information Forensics and Security (WIFS), (Tenerife, Spain), December 2012