IBM Research Report

The Relation between the ROC Curve and the CMC

R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, A. W. Senior
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598

IBM

The Relation Between the ROC Curve and the CMC

R.M. Bolle, J.H. Connell, S. Pankanti, N.K. Ratha and A.W. Senior
IBM Thomas J. Watson Research Center, Hawthorne, NY 10532
Exploratory Computer Vision Group
{bolle, jconnell, sharat, ratha, aws}@us.ibm.com

Abstract

The Cumulative Match Curve (CMC) is used as a measure of $1{:}\,m$ identification system performance. It judges the ranking capabilities of an identification system. The Receiver Operating Characteristic curve (ROC curve) of a verification system, on the other hand, expresses the quality of a 1:1 matcher. The ROC plots the False Accept Rate (FAR) of a 1:1 matcher versus the False Reject Rate (FRR) of the matcher. We show that the CMC is also related to the FAR and FRR of a 1:1 matcher, i.e., the matcher that is used to rank the candidates by sorting the scores. This has as a consequence that when a 1:1 matcher is used for identification, that is, for sorting match scores from high to low, the CMC does not *offer any additional information beyond the FAR and FRR curves. The CMC is just another way of displaying the data and can be computed from the FAR and FRR.*

1 Introduction

Performance evaluation of 1:1 biometric verification systems is achieved through estimating the False Accept Rate (FAR) and the False Reject Rate (FRR) and using these estimates to construct a ROC (Receiver Operating Characteristic) curve that expresses the tradeoff between the FAR and FRR [1].

The ROC is a well-accepted measure to express the performance of 1:1 matchers. How to measure the performance of identification systems is perhaps less clear. One measure is the Cumulative Match Curve (CMC) [2, 3, 4], which expresses the performance of biometric identification systems ($1{:}\,m$ search engines) that return ranked lists of candidates. Specifically, when an identification system always returns the identities associated with the K highest-scoring biometric samples from an enrollment database (gallery). To estimate the CMC, the match scores between a query sample and the m biometric samples in the database are sorted. The lower the rank of the genuine matching biometric in the en-

rollment database, the better the $1;m$ identification system.

The matcher is tested on multiple galleries and typically it is assumed that the same 1:1 matcher can be used for all galleries. The database (gallery) is an integral part of a $1{:}\,m$ search engine because without an enrollment database there is no identification system. However, the testing methodology of identification systems does not allow for tuning the system to the database.

Moreover, we show that there is a relationship between the ROC and the CMC associated with a 1:1 matcher (see also [1]). That is, given the ROC or given the FAR/FRR of a 1:1 matcher, the CMC can be computed and expresses how good this particular 1:1 matcher is at sorting galleries with respect to input query samples.

This paper is organized as follows. The next section introduces the CMC and Section 3 shows how to estimate a CMC given data samples. Section 4 introduces the closed world assumption under which the CMC is often estimated. In Section 5 the relation between the ROC and CMC is derived. Section 6 gives some empirical results and Section 7 ends with some conclusions.

2 The CMC

Assume that we have a (large) set of biometric samples B_i with associated ground truth $\text{ID}(B_i)$. Key to measuring a CMC curve associated with a 1:1 matcher is the assembly of two subsets of samples [4]:

1. A *gallery* set $\mathcal{G}$. There are m biometric samples in gallery set $\mathcal{G} = \{B_1, B_2, ..., B_m\}$; these are m biometric identifiers of *different* subjects.

   The gallery set $\mathcal{G}$ can be thought of as the enrolled database of biometric identifiers.

2. A *probe* set denoted as $\mathcal{Q}$ is a set of n "unknown" samples $(B'_1, B'_2, ..., B'_n)$ or $\{B'_\ell, \ell = 1, ..., n\}$, associated with the n subjects.

   The probe or query set $\mathcal{Q}$ can be from any set of individuals. However, usually probe identities are pre-

sumed to be in the gallery $\mathcal{G}$. The probe set may contain more than one biometric sample of a given person and need not contain a sample of each subject in $\mathcal{G}$.

Given a query biometric $B'_\ell \in \mathcal{Q}$ and a biometric $B_i \in \mathcal{G}$, the output of a biometric matcher is a similarity score $s(B'_\ell, B_i)$. In order to estimate the Cumulative Match Curve, each probe biometric is matched to every gallery biometric and a total of $n \times m$ similarity scores is computed,

$$\mathcal{S}_\ell = \{s(B'_\ell, B_1), s(B'_\ell, B_2), ..., s(B'_\ell, B_m)\}, \ \ell = 1, ..., n.$$

We have n sets $\mathcal{S}_\ell$ of m similarity scores each. The scores $s(B'_\ell, B_i), i = 1, .., m$ for each probe biometric $B'_\ell, \ell = 1, ..., n$ are ordered as

$$s(B'_\ell, B_{(1)}) \geq s(B'_\ell, B_{(2)}) \geq ... \geq s(B'_\ell, B_{(m)}) \qquad (1)$$

and probe B'_ℓ is assigned the rank $k_\ell = k$ if the matching sample from $\mathcal{G}$ is $B_{(k)}$. Hence the rank estimate k_ℓ of probe B'_ℓ is k if the matching gallery biometric is in the k-th location of the sorted list of (1), or equivalently, if the matching identity is in the k-th location of the associated list (or vector) of biometric identifiers,

$$\mathbf{C}_m(B'_\ell; \mathcal{G}) = (B_{(1)}, B_{(2)}, ..., B_{(k)}, ..., B_{(m)}). \qquad (2)$$

We denote this list as $\mathbf{C}_m(B'_\ell; \mathcal{G})$ because the sorted list depends on input query B'_ℓ and on the gallery $\mathcal{G}$.

3 Estimating the CMC

We have set $\mathbf{K}$ of n rank estimates $\{k_\ell; \ell = 1, ..., n\}$ (with $1 \leq k_\ell \leq m$), one estimate for each probe biometric B'_ℓ; each rank is defined *only if* the correct identity is in the ordered list of gallery biometrics $\mathbf{C}_m(B'_\ell; \mathcal{G})$ of (2).

Before going into the Cumulative Match Curve (CMC), let us first define the discrete rank probabilities $P(k), k = 1, ..., m$, of a biometric search engine. These probabilities associated with a search engine are simply the probabilities, summing to 1, that the identity associated with a probe has rank k. The $P(k)$ are the true frequencies of occurrence, or discrete probabilities, of ranks $1 \leq k \leq m$. Basically, $P(k)$ could be any discrete Rank Probability Mass (RPM) function. However, a Rank Probability Mass function with low average rank, is preferred.

Given the probe and gallery data, the $P(k)$ associated with a biometric matcher, are estimated by

$$\hat{P}(k) = \frac{1}{n}(\# \, k_\ell = k) \qquad (3)$$

$$= \frac{1}{n}(\# \, k_\ell \in \mathbf{K} = k), \ \ k = 1, ..., m.$$

The probability $\hat{P}(k)$ that a matching biometric has rank k is estimated as the fraction of probe biometrics B'_ℓ for which

$k_\ell = k$; $\hat{P}(1)$ is an estimate of the probability that the rank of any probe is 1, $\hat{P}(2)$ is an estimate of the probability that the rank is 2, and so on.

The Cumulative Match Curve estimates the distribution of the ranks $k_\ell, \ell = 1, ..., n$ of probes $\{B'_1, B'_2, ..., B'_n\}$; $\mathrm{CMC}(k)$ is the fraction of probe biometrics B'_ℓ that have rank $k_\ell \leq k$. That is,

$$\mathrm{CMC}(k) = \frac{1}{n}(\# \, k_\ell \leq k) \qquad (4)$$

$$= \frac{1}{n}(\# \, k_\ell \in \mathbf{K} \leq k) = \frac{1}{n}\sum_{\ell=1}^{n}\mathbf{1}(k_\ell \leq k).$$

Hence by definition, the true CMC is the sum of the $P(k)$

$$\mathrm{CMC}(k) = \sum_{r=1}^{k} P(r); \ \ k = 1, ..., m, \qquad (5)$$

(see Figure 1). The CMC estimate is the distribution of the estimated ranks, denoted as k_ℓ and estimates the probability $Prob(k_\ell \leq k)$. The random variable k_ℓ takes on discrete values $1, 2, ..., m$, determined by gallery $\mathcal{G}$ size m. Figure 1 shows a Cumulative Match Curve $\mathrm{CMC}(k)$; a higher $\mathrm{CMC}(1)$ and a steeper slope of the curve at $k = 1$ indicate a better 1:1 matcher to implement $1{:}m$ search through sorting.
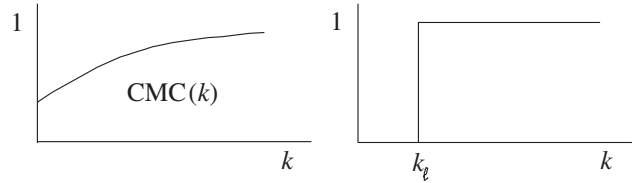

Figure 1. The CMC (left) is a weighted sum of step functions $\mathbf{1}(k_\ell \leq k)$ at $k = k_\ell$ as on the right

To estimate (4) probes can be constructed in two ways: one of which is with the *closed universe assumption*. Every probe biometric identity $\mathrm{ID}(B'_\ell)$ is some identity $\mathrm{ID}(B_i)$ in the gallery $\mathcal{G}$ of size m.

4 Closed universe

Each probe biometric B'_ℓ has a corresponding gallery biometric B_i and the correct match is always somewhere in the return vector of (2). This means that for each probe $\ell = 1, ..., n$, a rank k_ℓ, $1 \leq k_\ell \leq m$, is defined and can be determined from reordered list

$$\mathbf{C}_m(B'_\ell; \mathcal{G}) = (B_{(1)}, B_{(2)}, ..., B_{(m)}). \qquad (6)$$

To identify an unknown sample B'_ℓ from the samples $B_i, i = 1, ..., m$ in gallery $\mathcal{G}$, let the decision rule be to choose the highest-scoring candidate in vector $\mathbf{C}_m(B'_\ell; \mathcal{G})$ of (6) as the correct answer:

$$\text{Decide ID}(B'_\ell) = \text{ID}(d_{(1)}), \qquad (7)$$

the top sample from reordered vector $\mathbf{C}_m(B'_\ell; \mathcal{G}) = (B_{(1)}, B_{(2)}, ..., B_{(m)})$.

Selecting the top gallery biometric $\mathbf{C}_1(B'_\ell; \mathcal{G}) = B_{(1)} \in \mathcal{G}$ as the correct match may appear to be a good identification strategy, because identifier B'_ℓ is in gallery $\mathcal{G}$ and may be believed to be the most likely candidate on the top of the ranked list. The estimated rank probability $\hat{P}(k = 1)$ of (3) is an estimate of the probability of correct identification, under this assumption. A subject cannot be falsely rejected and cannot be falsely accepted; he or she can *only* be misidentified.

Alternatively, to make a decision, one can construct a short vector (candidate list) $\mathbf{C}_K(B'_\ell; \mathcal{G})$ of identities of the K top-scoring gallery biometrics B_i. This list of identities associated with the top K gallery biometrics, $\mathbf{C}_K(B'_\ell; \mathcal{G}) = (B_{(1)}, B_{(2)}, ..., B_{(K)})^T$ comprises the K most likely identities.

Now, suppose the gallery $\mathcal{G}$ is some "most wanted" list and the decision rule is

$$
\begin{aligned}
&\text{ID}(B'_\ell) \text{ is "Most wanted,"} && \text{if } B'_\ell \in \mathbf{C}_K(B'_\ell; \mathcal{G}), \text{ or,} \\
&\text{ID}(B'_\ell) \text{ is "Most wanted"} && \text{if rank of } B'_\ell \leq K.
\end{aligned}
$$
$$(8)$$

That is, a correct decision is deemed to have been made if the true identity of unknown sample B'_ℓ is in the candidate list of length K. This is the *negative identification* problem: "Is subject $\text{ID}(B'_\ell)$ em not on some list?" Of course, the closed world assumes that *every* subject $\text{ID}(B'_\ell)$ that is authenticated is on this "wanted" list.

This assumption will be relaxed in Section 5.3.

As is the case when we only consider the top candidate on the list (7), also for $K > 1$, a sample B'_ℓ cannot be falsely rejected and cannot be falsely accepted, he or she can only be mistakenly left off the candidate list. Again, only identification errors can be made, with probability

$$Prob\,(B'_\ell \text{ is misidentified; } K) = \hat{P}_F(K). \qquad (9)$$

This error probability depends first of all on candidate list size K and is lower for larger K. It further depends on the gallery size m and is higher for larger m. A third influence is the quality of the match engine, or the distributions of genuine and imposter scores. The probability of correct identification then is

$$Prob\,(B'_\ell \text{ is identified; } K) = \text{CMC}(K) = 1 - \hat{P}_F(K).$$

The Cumulative Match Curve is an estimate of the probability of correct identification as a function of candidate list size K identification. The discrete function $\text{CMC}(K)$ converges to 1 when K approaches m. The faster the convergence, the more likely it is that the genuine matches are in short candidate lists for a given biometric matcher.

5 The CMC, FAR/FRR relation

Even though the decision rule of (8) does not use any thresholds, the match engine produces scores $s(B'_\ell, B_i)$, and we can relate the CMC to the False Accept and False Reject rates of the underlying 1:1 biometric matcher used in the $1\!:\!m$ search engine, by realizing that for every search the genuine score of the correct template can be interpreted as a "virtual threshold" on the imposter scores $s(B'_\ell, B_i)$, $\text{ID}(B'_\ell) \neq \text{ID}(B_i)$ from the rest of the database. (Note that this relation between the CMC and the FAR/FRR only holds when the 1:1 matcher is used for sorting the gallery scores.)

Each probe $\mathcal{Q}$ is a biometric B'_ℓ with a corresponding biometric B_i in gallery $\mathcal{G}$. Matching probe biometric B'_ℓ to a gallery of size m is equivalent to drawing of m random variables:

1. One genuine score $X = s(B'_\ell, B_i)$ for $B_i \in \mathcal{G}$ with B'_ℓ and B_i from the same identity.

2. A set of $(m - 1)$ imposter scores $\{Y_1, ..., Y_{m-1}\}$, i.e., $\{s(B'_\ell, B_i), i = 1, ..., (\ell - 1), (\ell + 1), ..., m\}$, where the B_i are in the gallery but not including biometric B_i matching to B'_ℓ.

To look at this, we turn to the rank probabilities of (3) first to simplify the analysis

$$\hat{P}(k) = Prob\,(\text{rank } B'_\ell \text{ is } k), \quad k = 1, ..., m.$$

These can be determined using joint probabilities $p(x, y_1, ..., y_{m-1})$ of the m random variables $\{s(B'_\ell, B_i), i = 1, ..., m\}$.

5.1 Gallery size $m = 2$

Pick a gallery size of $m = 2$ for the moment; this can be depicted in the xy plane. Any probe biometric B'_ℓ with $\ell = 1$ will result in a pair of random variables (X, Y), with $X = s(B'_\ell, B_1)$ a genuine score and $Y = s(B'_\ell, B_2)$ an imposter score. The genuine score is from probability density function (PDF) $f(x)$; the imposter scores are from PDF $g(y)$, each associated with the 1:1 matcher. There exist only two rank probabilities

$$
\begin{aligned}
P(1) &= Prob\,(\text{rank } B'_\ell \text{ is } 1) = Prob\,(X \geq Y), \\
P(2) &= Prob\,(\text{rank } B'_\ell \text{ is } 2) = Prob\,(Y > X).
\end{aligned}
$$

Let the joint probability density for a probe (X, Y) be $p(x, y) = f(x)\,g(y)$, assuming independence of the genuine and imposter score (see Figure 2). The rank probabil-
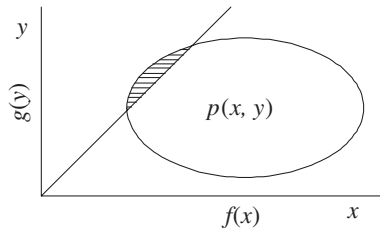
Figure 2. The probability of misidentification when gallery size is $m = 2$

Figure 3. For gallery size $m = 2$, the probability of correct identification is the integral of $f(x)[1 - \mathrm{FAR}(x)]$ and $g(y)[1 - \mathrm{FRR}(y)]$ as in (10)

ity $P(1)$ is the integral of $p(x, y)$ over the area $x > y$ and is the probability of correct identification. Rank probability $P(2) = 1 - P(1)$, the probability of misidentification, is the integral of $p(x, y)$ over the area $x < y$. These two rank probabilities can be written down.

First look at the probability of correct identification, $P(1)$. In Figure 2, this is the probability that genuine score $X > Y$. This is the probability mass for $x > y$

$$
\begin{aligned}
P(1) &= Prob\,(X > Y) = \int\!\!\int_{x>y} p(x, y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{\infty} f(x) \int_{y=0}^{x} g(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{y=0}^{\infty} g(y) \int_{x=y}^{\infty} f(x)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{x=0}^{\infty} f(x)\,[1 - \mathrm{FAR}(x)]\,\mathrm{d}x \\
&= \int_{y=0}^{\infty} g(y)\,[1 - \mathrm{FRR}(y)]\,\mathrm{d}y \qquad (10)
\end{aligned}
$$

The interpretation of (10) is shown in Figure 3. Given any genuine score $X = x$, the probability that $y < x$ is the probability that y is not a False Accept when adjusting threshold t_o to declare score x a match, i.e., is $[1 - \mathrm{FAR}(x)]$. Conversely, given any y, the probability that $x > y$ is $[1 - \mathrm{FRR}(y)]$, the probability that x is not a False Reject at threshold y.

What is the probability that the correct identity has rank 2, i.e., what is $P(2)$? From Figure 2 it is seen that the probability that the imposter similarity score Y is larger than X is the integral of $p(x, y)$ over the shaded region $y > x$

$$
\begin{aligned}
P(2) = 1 &- P(1) = Prob\,(X < Y) \\
= \int\!\!\int_{x<y} p(x, y)\,\mathrm{d}y\,\mathrm{d}x &= \int_{x=0}^{\infty} f(x)\,\mathrm{FAR}(x)\,\mathrm{d}x \\
&= \int_{y=0}^{\infty} g(y)\,\mathrm{FRR}(y)\,\mathrm{d}y. \qquad (11)
\end{aligned}
$$

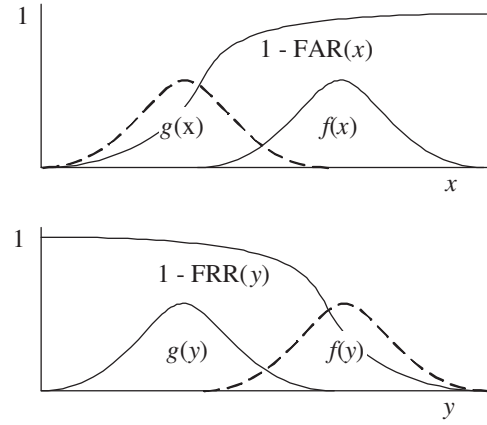Interpretations of these probabilities are as above. The first integral of (11) is shown in Figure 4a. Given some genuine score $X = x$, the probability that an imposter score y is larger than x is $\mathrm{FAR}(x)$, the probability of a False Accept at threshold x. The shaded area in the detail of Figure 4b is the probability of incorrect identification; this is the integral $f(x)\,\mathrm{FAR}(x)$ over all possible genuine scores x.

The second integral of (11) is an integral of possible values for imposter scores y. For any y the probability that genuine score x is less than y is the probability of a False Reject at threshold y; the integral over $g(y)\,\mathrm{FRR}(y) = P(2)$.

5.2 Gallery size $m > 2$

When $m > 2$, a probe $\mathcal{Q}$ is the drawing of *one* random genuine score X and the drawing of $(m - 1)$ random imposter scores Y_i, $i = 1, ..., m - 1$.

Let us first examine rank probability $k = 1$, i.e., $P(1)$. This is $Prob\,(\text{rank } X \text{ is } 1)$, hence the probability that $X > Y_1, X > Y_2, ..., X > Y_{m-1}$. Again assuming independence, the joint probability density of $(x, y_1, y_2, ..)$ is

$$
p(x, y_1, ..., y_{m-1}) = f(x) \prod_{i=1}^{m-1} g(y_i).
$$

The rank probability $P(1)$ is the integral of $p(x, y_1, y_2, ..)$, as in Figure 2, but over a hyper region this time:

$$
\begin{aligned}
P(1) &= \\
\int_x \int_{y_1 < x} &\cdots \int_{y_{m-1} < x} p(x, y_1, ..., y_{m-1})\,\mathrm{d}x\,\mathrm{d}y_1 \cdots \mathrm{d}y_{m-1} \\
&= \int_{x=0}^{\infty} f(x) \left(\prod_{i=1}^{m-1} \int_{y=0}^{x} g(y_i)\,\mathrm{d}y_i \right) \mathrm{d}x.
\end{aligned}
$$

$$(12)$$

FAR (x)

$g(x)$ $f(x)$

x

a

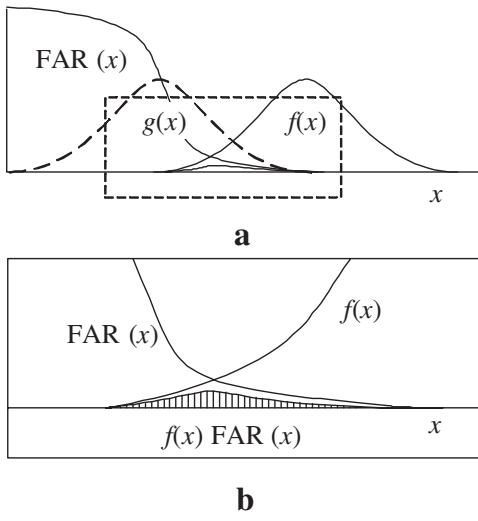FAR (x) $f(x)$

$f(x)$ FAR (x) x

b

Figure 4. For gallery size $m = 2$, the probability of misidentification is the integral of $f(x)\,\mathrm{FAR}(x)$

Using (10), this becomes

$$P(1) = \int_{x=0}^{\infty} f(x)\,[1 - \mathrm{FAR}(x)]^{m-1}\,\mathrm{d}x. \qquad (13)$$

Figure 5 compares $P(1)$ for $m = 2$ of (10) with the $P(1)$ for $m > 2$ of (13). The probability $[1 - FAR(x)]^{m-1} \ll [1 - FAR(x)]$ for large m and $Prob$ (rank X is 1) obviously becomes less and less for larger m.

1

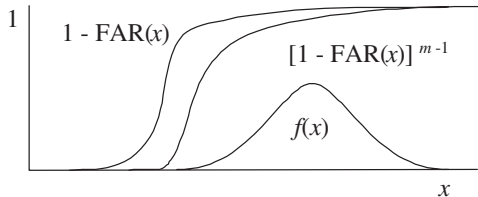1 - FAR(x)

$[1 - \mathrm{FAR}(x)]^{m-1}$

$f(x)$

x

Figure 5. The probability $P(1)$ that genuine score X is assigned rank 1 goes down for increasing m

By extending the above, we are now ready to construct $P(k)$, the probability that the "some" correct identity ends up in the k-th position of the return vector $\mathbf{C}_m$. The probability that rank X is k is the probability that $(k - 1)$ of the $(m - 1)$ imposter scores $Y_1, ..., Y_{m-1}$ are greater than X.

This is given by

$$P(k) = \binom{m-1}{k-1} \times$$
$$\int_x \int_{y_1 \cdots y_{k-1} > x} \int_{y_k \cdots y_{m-1} < x} p(x, ..., y_{m-1})\, \mathrm{d}x \cdots \mathrm{d}y_{m-1},$$

accounting for the number of ways $(k - 1)$ imposter scores can be selected from $(m - 1)$. Using (10) and (11), this becomes

$$\binom{m-1}{k-1} \int_0^\infty \left(\prod_{i=1}^{k-1} \int_{y=x}^\infty g(y_i)\,\mathrm{d}y_i\right) f(x) \times$$
$$\left(\prod_{i=k}^{m-1} \int_{y=0}^x g(y_i)\,\mathrm{d}y_i\right) \mathrm{d}x.$$

And we have

$$P(k) =$$
$$\binom{m-1}{k-1} \int_0^\infty [\mathrm{FAR}(x)]^{k-1} f(x)\,[1 - \mathrm{FAR}(x)]^{m-k}\,\mathrm{d}x. \qquad (14)$$

The interpretation of (14) is as above. Given some genuine score $X = x$ with probability density $f(x)$, there are $(k - 1)$ imposter scores Y_ℓ greater than x with probability $[\mathrm{FAR}(x)]^{k-1}$. There are an additional $(m - k)$ imposter scores Y_ℓ that are less than x with probability $[1 - \mathrm{FAR}(x)]^{k-1}$.

An approximate interpretation of (14) is the following. If we assume that the genuine score PDF $f(x)$ is relatively narrow and can be well approximated by an impulse function $\delta(x - \hat{x})$, with $\hat{x}$ the expected match score of B'_ℓ, (14) becomes

$$P(k) \approx \binom{m-1}{k-1} [\mathrm{FAR}(\hat{x})]^{k-1}\,[1 - \mathrm{FAR}(\hat{x})]^{m-k}. \qquad (15)$$

This is the probability that $(k - 1)$ imposter scores Y are greater than $\hat{x}$, *and* $(m - k)$ imposter scores Y are less than $\hat{x}$.

Remember, the Cumulative Match Curve $\mathrm{CMC}(K)$ of (17) is an estimate of the probability of correct identification as a function of candidate vector $\mathbf{C}_K$ size K. A sample B' is "correctly identified" if the corresponding identity in the gallery has rank K or less, and following (5), we get for the true CMC

$$\mathrm{CMC}(K) = \sum_{k=1}^K P(k) =$$
$$\sum_{k=1}^K \binom{m-1}{k-1} \int_0^\infty [\mathrm{FAR}(x)]^{k-1} f(x)\,[1 - \mathrm{FAR}(x)]^{m-k}\,\mathrm{d}x. \qquad (16)$$

When using (15), this becomes approximately

$$\mathrm{CMC}(K) \approx$$
$$\sum_{k=1}^K \binom{m-1}{k-1} [\mathrm{FAR}(\hat{x})]^{k-1}\,[1 - \mathrm{FAR}(\hat{x})]^{m-k}, \qquad (17)$$
$$K = 1, ..., m.$$

The probability that the true rank k of the correct identity $\leq K$ is the sum of the true individual probabilities that rank is k from $k = 1$ to $k = K$.

One thing about Expression (17) between the CMC of a $1{:}\,m$ search engine and the FAR/FRR of a 1:1 match engine should be noted again. Given the way the CMC is computed by *sorting* $s(B'_\ell, B_i)$ as in (1) is only *one* particular way of implementing $1{:}\,m$ search based an a 1:1 search engine.

5.3 Open world

Suppose there are an additional $n_{unknown}$ biometrics of unknown samples B' in the probe. Then the CMC is weighted by the prior probability that $B' \in \mathcal{G}$

$$\frac{n_{known}}{n_{known} + n_{unknown}} = \frac{m}{m + n_{unknown}}.$$

While the CMC is a useful characterization of a system in controlled evaluations, it is only a partial characterization of the performance of a system.

6 Experiment

Figure 6 shows a plot of an empirical CMC versus its theoretical form. The theoretical form was estimated by first computing the FAR and FRR distributions from the matcher scores and then using Expression (16) to estimate $\mathrm{CMC}(K)$ for K ranging from 1 to the size of the gallery. Given the matcher scores, the empirical form was estimated by calculating the CMC directly from its definition and without first computing the FAR or FRR distribution.
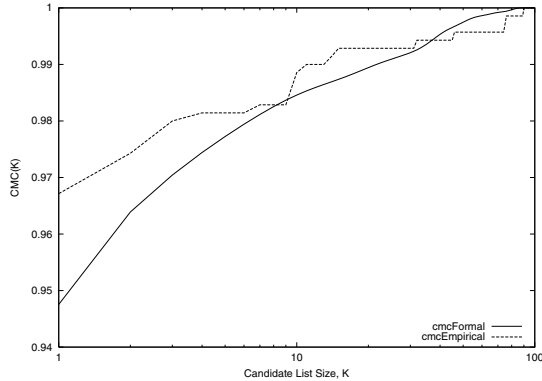

Figure 6. Theoretical versus empirical form of the CMC.

It is easy to see that as candidate list size, K increases, the error decreases and when K reaches the gallery size, both empirical and theoretical estimates (correctly) reach 1. For this very same reason, note that the difference in empirical and theoretical estimates decreases as the candidate list size increases. Further, it can be observed that the maximum discrepancy between theoretical and empirical estimates is

about two percent and it is empirical evidence supporting the integrity of our theoretical derivation.

7 Conclusions

Performance evaluation of biometric systems is an important topic. The generally accepted method for expressing the performance of a 1:1 verification system is an ROC curve. Typically, the performance of a $1{:}\,m$ identification system is expressed by a Cumulative Match Curve (CMC), which is really a discrete function. In this paper we have shown that the CMC is directly related to the ROC, the tradeoff between the FAR and FRR as a function of the operating threshold. That is, given the characteristics of a 1:1 matcher in terms of genuine and imposter score distributions or estimates thereof, the CMC can be constructed. Hence, as it is used today, the CMC is more a performance measure of a 1:1 matcher (verification system) than a performance measure of a $1{:}\,m$ identification system.

When computed with the "probe-gallery" approach, the CMC merely expresses the sorting capabilities of a 1:1 matcher. A better way of testing $1{:}\,m$ identification systems is to make the galleries public some time before the test and sequester the probe biometric samples till the time of the test. It is then up to identification system designer how to use a 1:1 matcher in the identification task—if a 1:1 matcher is used at all.

In our experiment, we have used a very simple identification scheme based on a 1:1 matcher. Note that in many applications, the identification process may involve more than a matcher (see, for example, [1], Chapter 16). In such situations, the discrepancy between empirical CMC (from identification procedure) and the theoretical CMC (from theoretical analysis of 1:1 FAR and FRR using Expression (16) can be an important metric for assessing the additional value provided by identification process.

References

[1] R. Bolle, J. Connell, S. Pankanti, N. Ratha, and A. Senior. *Guide to Biometrics*. Springer, New York, 2004.

[2] P. Grother and P. Philips. Models of large population recognition performance. In *Proc. CVPR*, volume 2, pages 68–75, 2004.

[3] P. Phillips, P. Grother, R. Micheals, D. Blackburn, T. Elham, and J. M. Bone. FRVT 2002: Facial recognition vendor test. Technical report, DoD, Apr. 2003.

[4] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE PAMI*, 22(10):1090–1104, Oct. 2000.