Studies of Plain-to-Rolled Fingerprint Matching Using the NIST Algorithmic Test Bed (ATB)

NISTIR 7112

Stephen S. Wood
Charles L. Wilson

April 2004

Studies of Plain-to-Rolled Fingerprint Matching Using the NIST Algorithmic Test Bed (ATB)

Stephen S. Wood & Charles L. Wilson
(from the National Institute of Standards & Technology)

EXECUTIVE SUMMARY

A series of fingerprint matching studies have been conducted on an experimental laboratory system called the Algorithmic Test Bed (ATB). The ATB is a system used to test the functional characteristics of the Automated Fingerprint Identification System (AFIS) component of the Federal Bureau of Investigation's (FBI) Integrated Automated Fingerprint Identification System (IAFIS). It was supplied by Lockheed Martin[a], the vendor of IAFIS, whose technology it shares. The ATB includes a built-in gallery of nearly 1.2M subjects, effectively a 3% representative sample of the FBI's criminal master file, with some room for additions. The ATB provides broad control over its operating modes and set points.

The ATB, like IAFIS, was originally designed to identify (and verify) subjects using submitted (predominantly rolled) fingerprint cards. The database against which the submitted cards were to be matched was likewise predominantly composed of rolled images from fingerprint cards. As demand rose for the use of plain (or flat or slap) images, ascertaining the operational characteristics of this matcher with plain submissions became important.

These studies were organized in three major groups. Studies of plain-to-rolled matching as a function of data source made up the first group. Studies of plain-to-rolled matching using large sets of operational data made up the second. Studies of plain-to-plain matching made up the third.

The studies answered nine major questions:

(1) how accurate is plain-to-rolled identification? is it less accurate than rolled-to-rolled?
Failure to identify (1-TAR) can be 2% or less. Accuracy can be as good as with rolled prints of average quality.

(2) how accurate is plain-to-rolled identification with fewer than ten fingers?
Accuracy with fewer than ten fingers can exceed accuracy with ten fingers, but performance penalties can be expected (see next question).

(3) are performance penalties incurred by adopting plain-to-rolled identification? if penalties are incurred, how great are they?
Performance (throughput) deteriorates markedly as the amount of data is reduced. Using plain rather than rolled fingerprints in 10-finger or 8-finger matching can be expected to result in a three-fold reduction in throughput. Reducing the number of fingers from ten to two can be expected to reduce throughput by a factor of ten or more. These effects are independent of one another, and thus multiplicative.

---

[a] *Specific hardware and software products identified in this report were used in order to adequately support the development of technology to conduct the performance evaluations described in this document. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose.*

Studies of Plain-to-Rolled Fingerprint Matching

(4) are there significant differences in accuracy and performance between the fingerprint images supplied by two different models of livescan devices? if there are differences, how great are they?

There are significant differences in both matching accuracy and performance. In this case, the inferior set of images showed a rate of failure to identify (1-TAR) 68% to 94% higher than that shown by the superior, and a filter rate 24% to 60% higher.

(5) does gallery size affect accuracy? if there is an effect, how can it be characterized?

Within the range of gallery sizes covered by these studies (160,000 to 42,000,000), gallery size did not affect overall accuracy (TAR).

(6) does the ATB accurately model the behavior of IAFIS?

In a word, yes.

(7) how accurate is plain-to-plain identification?

Failure to identify (1-TAR) can be 2% or less. There is more variability than evident with rolled-to-plain.

(8) how accurate is plain-to-plain identification with fewer than ten fingers?

As with rolled-to-plain, accuracy with fewer than ten fingers can exceed accuracy with ten fingers, with a failure rate of well under 1%. There is more variability than evident with rolled-to-plain, and performance penalties can be expected.

(9) are performance penalties incurred by adopting plain-to-plain identification?

Yes. Filter rates can approach 50%, which means matching every other subject. (Compared to the 2% filter rate typical of rolled-to-rolled tenprint matching, a 50% rate implies a 25-fold increase in matching workload.)

We were able to draw other conclusions as well:

There were clear differences, but the interclass difference (between rolled and plain) was smaller than the intraclass differences (between BCII and Identix and between CrossMatch and Smiths Heimann).

There were similarities among all the sets; the invariance of TAR over gallery size and the essentially linear relationship of FAR to gallery size were both notable.

The ATB works most accurately (i.e., has the lowest failure rate, or (1-TAR)) with two index fingers. The addition of more fingers raises its throughput markedly but with a concomitant reduction in TAR. The use of six or more fingers reduces the FAR to almost zero.

Studies of Plain-to-Rolled Fingerprint Matching

1.   INTRODUCTION

NIST recently conducted a series of fingerprint matching studies using an experimental laboratory system called the Algorithmic Test Bed (ATB). NIST's specimen of the ATB, a system used to test the functional characteristics of the AFIS component of the FBI's Integrated Automated Fingerprint Identification System (IAFIS), includes a gallery of nearly 1.2M subjects; it provides broad control over its operating modes and set points [6]. This report focuses on one aspect of those studies: the matching of plain to rolled (and plain to plain) fingerprint images[b].

The importance of measuring the efficacy of matching plain to rolled and plain to plain images stems from (at least) three sources:

(1) the desirability of rapidly and expeditiously fingerprinting subjects without compromising the ability to compare the fingerprints thus captured to those vast sets of rolled fingerprints already in legacy databases: although capturing plain fingerprints is faster and easier than capturing rolled fingerprints, rolled fingerprints predominate in existing databases [5];

(2) the passage of PL 107-56 and PL 107-173, that effectively require NIST to determine the efficacy of both plain-to-rolled and plain-to-plain matching (as potential solutions to some of the many mandates of these laws) [1,2]; and

(3) the implicit requirements imposed by the Daubert decision on the criminal justice community for the statistical data necessary to provide the quantitative, scientific basis for the continued use of expert testimony in the field of fingerprint identification and verification [7,8,9].

There were two additional reasons for performing these studies:

(1) it is of value (particularly to both NIST and the FBI) to know whether the ATB closely models the IAFIS with respect to (w.r.t.) fingerprint matching, and therefore whether fingerprint matching experiments performed on the ATB can properly be used to make statements and predictions about IAFIS; thus we sought a calibration of the ATB, using the operational IAFIS as a baseline; and

(2) it is also of value (particularly to both NIST and the Ohio BCII) to know whether different mechanisms of acquiring images produce images of different matchability; thus we made a comparison, w.r.t. the accuracy of fingerprint matching, of images from different sources, acquired by different means.

Six sets of data were used, each from a different source. These six sets were respectively referred to as:

---

[b] Although we speak of matching images, the actual process of matching generally operates on 'features' or 'templates' which the matching system has extracted from the images in a preceding process.

Studies of Plain-to-Rolled Fingerprint Matching

     (1) BCII (Ohio Bureau of Criminal Identification and Investigation),
     (2) TXDPS (Texas Department of Public Safety),
     (3) INS10 ([former] Immigration and Naturalization Service (INS) tenprint),
     (4) INS_ben ([former] INS benefits),
     (5) INS_plain ([former] INS border apprehensions), and
     (6) BCC (Department of State (DoS) Border Crossing Cards).

These studies of this data fell into three broad groups.

The first group of studies, which looked at the effects of the source and type of data on plain-to-rolled matching, used only data from set 1, the Ohio BCII data.

The second group of studies, which examined the accuracy and throughput of plain-to-rolled matching using large sets of operational data, used only data from sets 2, 3, and 4, the Texas Department of Public Safety data, the INS tenprint data, and the INS benefits data, respectively.

The third group of studies, which explored the accuracy and throughput of plain-to-plain matching, used some data from set 1, together with data from two large sets, 5 and 6, respectively the INS border apprehensions and the DoS Border Crossing Cards.

Of these, four sets (1, 2, 3, and 5) had been captured as law enforcement data, while the remaining two sets (4 and 6) had been captured as civil administrative data.

The studies addressed nine questions pertinent to automated fingerprint identification.

    (1) how accurate is plain-to-rolled identification? is it less accurate than rolled-to-rolled?
    (2) how accurate is plain-to-rolled identification with fewer than ten fingers?
    (3) are performance [in the sense of speed or throughput] penalties incurred by adopting plain-to-rolled identification? if penalties are incurred, how great are they?
    (4) are there significant differences in accuracy and performance between the fingerprint images supplied by different models of livescan devices and different data sources? if there are, how great?
    (5) does gallery size affect accuracy? if there is an effect, how can it be characterized?
    (6) does the ATB accurately model the behavior of IAFIS?
    (7) how accurate is plain-to-plain identification?
    (8) how accurate is plain-to-plain identification with fewer than ten fingers?
    (9) are performance penalties incurred by adopting plain-to-plain identification?

These studies used the ATB in its normal operational mode: that is, filtering[c] aggressively [6].

---

[c] Filtering, together with several other terms, is described in the glossary (Appendix E). The term refers to reducing the number of potential candidates in order to reduce the number of matches to be evaluated. Because matching requires much more work than filtering, filtering aggressively increases throughput.

2. DATA

The characteristics of the six data sets used in this test are summarized below in table 1. A detailed description of each set is provided in the six subsections immediately following. The preparation of each data set is covered in section 3. Data analysis and the configuration of the matcher are described in section 4. Results are discussed in section 5. Conclusions are presented in section 6.

The first column of table 1 identifies the data set (or subset) by name and number. The second refers to its role in the plain-to-rolled matches: whether it provided members of the probe, the gallery[d], or both. The third column is similar, but refers to the plain-to-plain matches. The fourth column refers to the number of subjects whose images we used in these studies. The fifth column refers to the nature of the population from which the images were taken. The sixth shows the predominant[e] process by which the images were captured in digital form; the phrase 'mostly ink', in column six, refers to a set of images, most of which were digitized by optically scanning inked cards. The seventh describes the initial type of image; the word 'segment', in column seven, means that the images of the eight fingers were segmented from two slap-4 images by a computer code.

Data Set Number, Data Set Name	role: plain-rolled	role: plain-plain	working size	criminal civil	livescan rescan mostly ink	rolled plain segment
ATB background gallery	gallery		1.2M	criminal	mostly ink	rolled
1 Ohio BCII	-	-	-	-	-	-
1.1 BCII master	probe		925	criminal	mostly ink	rolled
1.2 Identix	probe		925	"	livescan	rolled
1.3 CrossMatch	probe	pr/ga	923	"	livescan	segment
1.4 SmithsHeimann	probe	pr/ga	925	"	livescan	segment
1.5 IAFIS master	gallery		925	"	mostly ink	rolled
1.6 BCII plain		pr/ga	630	"	mostly ink	segment
1.7 Identix plain		pr/ga	904	"	livescan	segment
1.8 IAFIS plain		pr/ga	793	"	mostly ink	segment
2 TXDPS [TXDPS]	probe		60K	criminal	mostly ink	segment
3 INS10 [DHS10]	probe		48K	criminal	rescan	segment
4 INS_ben	probe		60K	civil	livescan	segment
5 INS_plain [DHS2]		pr/ga	300K	criminal	livescan	plain
6 BCC [DOS]		pr/ga	274K	civil	livescan	plain

(Table 1   ATB data sets)

2.1   The Ohio BCII data (set 1) initially comprised five subsets of comparable images, each subset nominally containing a tenprint record for each of the 925 subjects involved in the study

---

[d] A probe is a set of records of nominally unknown subjects for which we seek mates in a set of known subjects, or gallery; see glossary.
[e] Subsets 1.1, 1.5, 1.6, and 1.8, together with sets 2, 3, and 4, all seem to have been acquired and digitized using more than one process; however, in each one process seems to have predominated.

[14]. The State of Ohio's Bureau of Criminal Identification and Investigation (BCII) provided the first four of the subsets on CDROMs; the Federal Bureau of Investigation's (FBI's) Criminal Justice Information Services (CJIS) Division similarly provided the fifth subset.

The same 925 subjects supplied the images within each subset, but a different method was used to acquire the images of each subset. Two of the subsets were acquired from criminal master files (CMF) from BCII and from the FBI, respectively; the other three were acquired by each of three models of livescan fingerprint machines (Identix TP600/2000, CrossMatch CMT 442, and Smiths Heimann LS2-Check RJ0444, respectively [14]). In both CMF subsets inked images predominated, although both subsets also contained substantial numbers of livescan images (see Appendix D). Of the three livescan machines, the Identix produced both rolled and plain impressions, whereas the other two produced only plain impressions.

The first subset (1.1, BCII) of data set 1 was taken from the BCII's master file. The second through fourth subsets (1.2, Identix; 1.3, CrossMatch; 1.4, Smiths Heimann) were acquired (apparently in the last week of June through the first week of July of 2003) by fingerprinting each subject on each of the three makes of livescan machines. CJIS provided the fifth subset of data (1.5, IAFIS), which it created by extracting the images for each of the 925 subjects from the master files of its Integrated Automated Fingerprint Identification System.

Three additional subsets of data were derived from set 1: 1.6, BCII_plain; 1.7, Identix_plain; and 1.8, IAFIS_plain. They are described in some detail under data preparation for plain-to-plain matching (section 3.3)

Early in the course of the experiment, the existence of three major difficulties with using the BCII subset to seed the gallery became apparent. First, images of three subjects were missing from the BCII subset, although these images were later supplied. Second, there were some discrepancies in the 'FBI number'[f] field in the cross-reference table. Third and most important, because the images in the FBI's master file were often not the same as those in the BCII's master file, comparability between the FBI's results and NIST's results would be compromised. In order to overcome these difficulties, the IAFIS subset was acquired and used to seed the gallery.

It is important to note that there were three fundamentally different types of sources of these subsets: CMF in rolled and plain tenprint card format (BCII, IAFIS), livescan in rolled and plain tenprint card format (Identix), and livescan in plain-only format (CrossMatch, SmithsHeimann).

Most of both the BCII and IAFIS subsets had been optically scanned from paper tenprint cards. An extremely common tenprint card, used by the FBI as well as much of the United States law enforcement community, the FD-249, appears in Figure 1. The FD-249 has distinct areas allocated for each of 14 original images: ten rolled images (in the boxes labeled with the finger

---

[f] FBI number is a number assigned by the FBI to identify a subject's records; see glossary

position[g] and name, for example '1. R. THUMB'), two plain thumb images, two slap-4 images. In these BCII and IAFIS cards, inked images predominated (Appendix D has a full discussion).

LEAVE BLANK	TYPE OR PRINT ALL INFORMATION IN BLACK		FBI	LEAVE BLANK
	LAST NAME NAM FIRST NAME MIDDLE NAME			

STATE USAGE

ALIASES · CONTRIBUTOR O R I

DATE OF BIRTH DOB
Month   Day   Year

SIGNATURE OF PERSON FINGERPRINTED

THIS DATA MAY BE COMPUTERIZED IN LOCAL, STATE AND NATIONAL FILES

DATE ARRESTED OR RECEIVED DOA   SEX   RACE   HGT.   WGT.   EYES   HAIR   PLACE OF BIRTH POB

DATE   SIGNATURE OF OFFICIAL TAKING FINGERPRINTS

YOUR NO. OCA

LEAVE BLANK

CHARGE

FBI NO. FBI

CLASS. _____

SID NO. SID

REF. _____

FINAL DISPOSITION

SOCIAL SECURITY NO. SOC

NCIC CLASS - FPC

CAUTION ☐

1. R. THUMB	2. R. INDEX	3. R. MIDDLE	4. R. RING	5. R. LITTLE
6. L. THUMB	7. L. INDEX	8. L. MIDDLE	9. L. RING	10. L. LITTLE

LEFT FOUR FINGERS TAKEN SIMULTANEOUSLY   L. THUMB   R. THUMB   RIGHT FOUR FINGERS TAKEN SIMULTANEOUSLY

(Figure 1   FBI Form FD-249) [4]

Visual inspection of a small sample of records showed a minor difference between the BCII and IAFIS records: substitution of livescan images for inked images had proceeded frequently on a

---

[g] 'finger position', by convention, is a number from 1 (right thumb) through 10 (left little); see glossary

finger-by-finger basis in the former but generally on a subject-by-subject basis in the latter.

The Identix subset consisted of livescan images in tenprint card format (i.e., with 14 original images, 10 rolled and four plain). These images had remained in digital (electronic) form since the time of their initial acquisition; they had **not** been rendered onto paper and then scanned.

The records in the CrossMatch and Smiths Heimann subsets each consisted of livescan images taken from three original images (right slap-4, thumb 'slap-2' (i.e., left thumb and right thumb captured simultaneously), left slap-4), segmented[h] automatically, and used to generate a record in tenprint card format.

There is no representation that the same livescan machine was used for all of the acquisitions by a given make; in the case of Identix, the records themselves indicate that two different (but substantially similar: TP-600 and TP2000) models were used. The CrossMatch machine(s) was identified as a model '442' (now supplanted by the '500'). The Smiths Heimann machine(s) was identified as a model LS2 Check.

2.2   TXDPS data

The TXDPS data (set 2) was obtained from the Texas Department of Public Safety (TXDPS), and consisted of images obtained by scanning conventional tenprint fingerprint cards from the TXDPS master file. Inked cards predominated. However, not all of the fingerprint cards used the same layout as the FD-249. Some used substantially different positions for the impressions of the fingers. It is evident that there were problems physically registering some of the cards in the optical scanner, and in some cases even in selecting a scanner format appropriate to the card. Despite these problems, most images seemed to be properly scanned and registered. Although the data set was considerably larger, the ATB study only addressed the first 60K (after consolidation[i]) [4]. The reason underlying the limitation to 60K is elaborated under Data Preparation (below).

2.3   INS10 data

The INS10 (also DHS10) data (set 3) was obtained from the Immigration and Naturalization Service (INS), which then became the Bureau of Immigration and Customs Enforcement (BICE), and is now part of the U. S. Department of Homeland Security (DHS). In other reports this data may also be referred to as DHS10 data. Like the BCII Identix data (subset 1.2), it for the most part originally consisted of livescan images in tenprint card format (i.e., with 14 original images,

---

[h] 'Segmentation' can refer in this context to two distinct processes. The first occurs when a multi-finger (e.g., slap-4) image is resolved into its constituent fingerprint images. The second occurs when a tenprint card (e.g., an FD-249) is optically scanned and segmented into its 14 constituent fingerprint image areas. We shall refer to the first as segmentation and to the second as scanning (or scanning and registration).

[i] 'Consolidation' in this case refers to elimination of additional images pertaining to a given subject; see glossary

10 rolled and four plain). Unlike the BCII Identix data, most of it was then printed onto paper cards and subsequently re-scanned prior to its receipt by NIST [4]. Although 54K records were initially available, after consolidation only 48K unique subjects remained.

2.4   INS_ben data

The INS_ben data (set 4) was obtained from the Immigration and Naturalization Service. It is referred to as INS benefits data. Like the BCII Identix data (subset 1.2), it for the most part originally consisted of 14 livescan images in tenprint card format, and remained in electronic form from its acquisition until its receipt by NIST. Together with the BCC data (set 6), it is civil in origin. As with the TXDPS data (set 2), the ATB study only addressed the first 60K (after consolidation).

2.5   INS_plain data

The INS_plain (also DHS2) data (set 5) was obtained from the INS IDENT (Automatic Biometric Fingerprint Identification) System. It was taken from INS (or DHS) recidivist cases: repeat 'contact' with border crossing offenders. These plain (flat) pairs of fingerprints of left and right index fingers were captured in the field by border patrol operations, for the most part in the American Southwest, using DFR-90 livescan units [4]. Although over 600K subjects were available, scheduling limitations and our experimental design constrained our probe set to 60K.

2.6   BCC data

The BCC (also DOS) data (set 6) was obtained from the U. S. Department of State. Like the INS_plain data set, it consists of plain pairs of fingerprints of left and right index fingers captured using DFR-90 livescan units. Unlike the INS_plain data set, the images were captured in an office environment (consular offices) as part of a civil procedure (issuance of Border Crossing Cards) [4]. About 274K subjects were available, but scheduling limitations and our experimental design constrained our probe set to 60K.

Studies of Plain-to-Rolled Fingerprint Matching

3.   DATA PREPARATION

In preparing data, there were two absolute limits that affected set size: the allocated size of each of the two ATB repositories. The repository with a built-in gallery of 1.2M rolled tenprint records had additional space for only 60K more. The entire capacity of the (initially) empty repository was 600K. These limits of 60K and 600K had no effect on small sets (set 1). They profoundly affected the choices made for sets 2, 4, 5, and 6.

3.1   Data Preparation - Group 1

Each subset of set 1 appeared in its own directory, and had its own numbering schema. For the four subsets provided by BCII, BCII also provided a cross-reference table showing the four corresponding file names for each of the 925 subjects. The FBI provided an equivalent cross-reference table. We used these records as originally provided after performing some simple manipulations to ensure consistent file names, to insert fields required by the EFTS specification [12], and so on; no modifications to the images themselves were required.

3.2   Data Preparation - Group 2

Sets 2, 3, and 4 were all prepared differently from set 1. In each of these sets there were no (intentional) correspondences of records from different sources. Rather, the two problems were segmentation and consolidation.

Sets 2, 3, and 4 each contained both rolled and plain images. To generate two records (one of rolled images, one of plain) for each subject, the slap-4 images were segmented. The only records used were those from which the segmentation software extracted eight images[j]. The rolled images generated the gallery subset of each set, while the plain generated the probe. This generation was performed for the NIST VTB studies [4] using NFIS software [3]. Records produced in the course of those studies were available to and used by the ATB studies.

The consolidation of all three sets is described in Appendix E.

At the end of segmentation and consolidation, we had extracted probe sets of 60K records from sets 2 and 4, and 48K records from set 3. These we randomized into 10 (or 8) subsets of 6K each in order to develop standard deviations (and thus a statistically valid measure of our confidence in our numerical results) for our mean TAR and FAR scores.

3.3   Data Preparation - Group 3

Group 3 contained five subsets of set 1, together with two large sets, 5 and 6.

---

[j] Were the segmentor to extract images of three or fewer fingers from a slap-4 image, we would not be sure **which** fingers they were.

3.3.1   Data preparation - group 3 - plain subsets of set 1

There were three subsets of set 1 whose tenprint card records originally had contained both rolled images and plain images: the BCII criminal master file records, the Identix livescan records, and the IAFIS criminal master file records. We had already created subsets 1.1, 1.2, and 1.5, respectively, from the rolled images for the group 1 studies. We now created three new subsets, 1.6 BCII plain, 1.7 Identix plain, and 1.8 IAFIS plain, from the plain images. We used these, together with subsets 1.3 and 1.4, for the group 3 studies.

To create these three new subsets, we used a NIST Fingerprint Image Software (NFIS) code to extract the two plain thumb and two slap-4 images from each tenprint record [3]. We decompressed the two slap-4 images using another code from NFIS, and converted them to a Sun raster format [11]. We segmented them on the ATB's HP n-class processor using a segmentor code provided by CJIS. The segmentor was not always able to identify and extract four single finger images from each slap-4 image, which is eight single finger images per subject. It was not successful in extracting all eight such images for 295 subjects in the case of the BCII master cards, for 21 subjects in the case of the Identix cards, and for 132 subjects in the case of the IAFIS master cards. We kept all records for which it did extract eight images, but we discarded all others. We recompressed all retained images at a low compression (near 5.), and then created an EFTS record from these eight recompressed images and the two still-compressed plain thumbs.

Although the decompression and recompression of an image can present difficulties, it was unavoidable: the ATB will not accept uncompressed images. Use of an unagressive setting for recompression served to minimize any possible damage from degradation of the image.

Thus we were able to create new plain subsets with 630 subjects from the BCII master tenprint cards, 904 from the Identix, and 793 from the IAFIS master. Existing plain subsets CrossMatch and Smiths Heimann held 923 and 925, respectively, for a total of 4175 plain records in the five plain subsets. The number of subjects **common** to any pairing of plain subsets ranged from 545 (BCII master – IAFIS master) to 923 (CrossMatch – Smiths Heimann).

3.3.2   Data preparation - group 3 - large sets

Sets 5 and 6 were all prepared similarly to sets 2, 3, and 4 (in group 2) w.r.t. both segmentation and consolidation.

Sets 5 and 6 presented their own problems. The ATB was designed as a rolled-to-rolled tenprint matcher, and was provided with a background gallery built from 1.2M rolled tenprints. Although it is perfectly feasible to conduct searches of fewer than ten fingers (one can specify which fingers are to be employed in the search), it is not even possible to conduct a search against a gallery of plain images when the entire gallery is consists of rolled images, and thus has

no plain images against which to search. It was therefore necessary to provide our own background gallery. The ATB came with an empty (reusable) repository with the capacity for 600K subjects, into which we loaded our background gallery of plain images.

The ATB requires a minimum of 5 seconds to characterize[k] a subject's record; it does not matter whether the record contains 1, 2, or 10 fingerprint images. To build a gallery of 600K subjects, each record containing the two index fingers, requires at a minimum 3M seconds (or 35 days) of characterization time, in addition to the time necessary to load the records into gallery. Loading the 600K characterized records into the gallery itself requires about one week. Additional time is required to characterize the probe set. Other ATB functions (e.g., searching, adding records to a gallery) cannot be carried on at the same time as characterization.

These scheduling constraints, coupled with externally mandated priorities, compelled that the gallery size for the initial plain-to-plain matching be much smaller than the 600K capacity. When the first 60K INS_plain records had been characterized (both probe set and gallery set), we performed a study using a 60K probe set and 60K gallery set. As machine time became available, we were able to characterize the entire 274K records in the BCC set, and perform a BCC study using a 60K probe set against a 274K gallery set. As more time became available, we were able to characterize 240K more INS_plain records, or 300K in all. We added these to the 274K BCC records, so that we would be able to conduct still more studies, searching each 60K probe set against the combined 574K gallery.

---

[k] characterization, as viewed by the ATB, means extraction of the information used to make matches; see glossary

4.   DATA ANALYSIS

For all three groups of studies, we used our own scoring codes rather than the ATB's code to score and analyze the raw results produced by the ATB.

4.1   Data Analysis - Group 1

In tenprint mode the ATB identified 24 records in the BCII data with candidate matches in the 1.2M background gallery. For each of these 24 records, we compared the FBI number of the candidate from the background gallery to the FBI number of the IAFIS image used to seed the gallery. In all cases except one, the FBI numbers corresponded. The FBI confirmed that the exceptional case was a consolidation: the same subject had appeared in its database twice, under two different FBI numbers. It further confirmed that the consolidation has since been rectified in its master files.

For each of the 24 FBI numbers described above, we replaced the FBI number reported with the number of the corresponding record in our subset 5 (and purged duplicated reports of a match) prior to scoring. That is, we created a virtual consolidated gallery G^C (see Appendix G) by ignoring any matches to elements of $(G - G^C)$.

4.2   Data Analysis - Group 2

For sets 2, 3, and 4, we had already checked for consolidations. However, as mentioned above, these data sets contained enough probe images that it was possible to obtain even more meaningful statistics. We subdivided each actual probe set into randomly selected subsets of 6K each (a special randomizing code was written), searched with each of these subsets, and then determined the mean and standard deviation of the measures of the accuracy (e.g., TAR, FAR) of these probes. This in turn quantified our confidence in these measures of accuracy.

4.3   Data Analysis - Group 3

Group 3 can be subdivided into the plain subsets of the Ohio BCII data and the two large sets of two-finger plain data, 5 and 6.

4.3.1   Data analysis - plain subsets of BCII

The subsets of set 1 with plain images (1.3, CrossMatch; 1.4, Smiths Heimann; 1.6, BCII_plain; 1.7, Identix_plain; and 1.8, IAFIS_plain) were analyzed in a somewhat different manner. In all, 20 searches were conducted. Each subset was in turn treated as a probe; that probe was then used to search against each of the remaining four subsets, which took turns serving as the gallery. For example, with 1.3 as the probe, a search was made with 1.4 as the gallery, then with 1.6 as the gallery, then with 1.7 as the gallery, and then with 1.8 as the gallery; next, with 1.4 as the probe, a search was made with 1.3 as the gallery, then with 1.6 as the gallery, then with 1.7 as the

gallery, and then with 1.8 as the gallery; and so on.

The analysis of the BCII plain-to-plain matching is covered in more detail in the discussion of results.

4.3.2   Data analysis - large sets of two-finger plain

Sets 5 and 6 were analyzed in much the same manner as sets 2, 3, and 4. However, for these two sets we were conducting plain-to-plain matching, so that searching against the 1.2M rolled background supplied by the ATB would have been meaningless. Instead we used two types of galleries. The first type was composed of images from the same source or origin as the relevant probe set. The second type was a composite: the union of the set 5 target[l] and the set 6 target. As with sets 2, 3, and 4, both automated consolidation and random partition of the 60K probe into subsets (or subprobes) of 6K each were employed.

4.4   Matcher Configuration

The ATB was used in the default configuration, as delivered by Lockheed Martin (LM). This configuration is essentially that used by IAFIS. Details of this configuration appear in Appendix A. When computing the true accept rate (TAR) we treated the ATB as if it only returned the rank-1 score, although when determining the false accept rate (FAR), we counted any reported mis-identification[m].

---

[l] a target is a set of fingerprint records of known subjects; it is a superset of a gallery; see glossary

[m] TAR and FAR are described in the glossary (qv); the choice of a pessimistic scoring methodology for TAR stems from the fact that the ATB models an operational matcher; in an operational matcher, the cost of a human review of a long list of candidates is not trivial: the ideal length of the list in the case of a match is 1, and in the case of a non-match is 0

5. RESULTS

The results for each group are discussed in turn.

5.1 Results from Group 1

Accuracy of plain-to-rolled identification

Table 2 shows that plain-to-rolled identification matching can be quite accurate. This accuracy is shown by a TAR of 97.5% using the Smiths Heimann plain images, the same as the mean TAR obtained using rolled (i.e., BCII and Identix) images. The Smiths Heimann plain subset produced 23 failures to identify in 925 cases, while the rolled Ohio BCII master file images yielded 27 failures and the rolled images from an Identix livescan resulted in 19 failures. Table 2 also shows that plain-to-rolled can be significantly less accurate than rolled-to-rolled identification matching: the results using images from the CrossMatch show a value of (1-TAR) that is 68% greater than the aggregate from rolled-to-rolled matching.

[10-FINGER vs IAFIS rolled]	TAR	1-TAR (FRR)	FAR	filter rate
BCII: rolled, mostly paper	97.1%	2.9%	0.0%	1.3%
Identix TP600/2000: rolled	97.9%	2.1%	0.0%	1.3%
CrossMatch 442: flat-4	95.8%	4.2%	0.0%	7.2%
SmithsHeimann LS2: flat-4	97.5%	2.5%	0.0%	4.5%

(Table 2   BCII series, matches incorporating both index fingers, 10-finger)

Accuracy of plain-to-rolled identification with fewer than ten fingers

Tables 3 and 4 show that plain-to-rolled identification matching with fewer than ten fingers can also be quite accurate. Table 3 shows that eight finger (i.e., excluding the thumbs) matching achieved a TAR of 98.2% using the Smiths Heimann plain images, the same as the mean TAR obtained using the equivalent rolled images. Table 4 shows that two finger matching achieved a TAR of 98.4% using the Smiths Heimann images. This latter is a respectable TAR but inferior to the mean TAR of 99.1% achieved using rolled images. As with table 2, the use of eight finger and two finger plain images from a different source can also achieve results of markedly higher inaccuracy (3.5% vs. 1.8% and 2.7% vs. 1.6%, respectively).

[8-FINGER (2-5,7-10) vs IAFIS rolled]	TAR	1-TAR (FRR)	FAR	filter rate
BCII: rolled, mostly paper	97.9%	2.1%	0.0%	3.9%
Identix TP600/2000: rolled	98.5%	1.5%	0.0%	3.9%
CrossMatch 442: flat-4	96.5%	3.5%	0.0%	15.5%
SmithsHeimann LS2: flat-4	98.2%	1.8%	0.0%	10.5%

(Table 3   BCII series, matches incorporating both index fingers, 8-finger)

Studies of Plain-to-Rolled Fingerprint Matching

[2-FINGER (2,7) vs IAFIS rolled]	TAR	1-TAR (FRR)	FAR	filter rate
BCII: rolled, mostly paper	98.7%	1.3%	1.2%	29.8%
Identix TP600/2000: rolled	99.5%	0.5%	0.4%	29.8%
CrossMatch 442: flat-4	97.3%	2.7%	0.8%	50.7%
SmithsHeimann LS2: flat-4	98.4%	1.6%	0.9%	40.9%

(Table 4   BCII series, matches incorporating both index fingers, 2-finger)

Tables 5 and 6 show the results of plain-to-rolled identification matching using the fingers of only one hand (the right). Table 5 shows that employing four fingers (i.e., the right hand excluding the thumb), the ATB achieved a TAR of 94.3% using the Smiths Heimann plain images and 90.9% using the CrossMatch. Both of these were markedly inferior to either of the rolled-to-rolled results, which had a mean TAR just under 97.5%. Table 6 shows that using two fingers (the right index and middle) gave TARs that were almost the same as with four fingers.

[4-FINGER (2-5) vs IAFIS rolled]	TAR	1-TAR (FRR)	FAR	filter rate
BCII: rolled, mostly paper	96.8%	3.2%	0.3%	16.6%
Identix TP600/2000: rolled	98.1%	1.9%	0.3%	16.6%
CrossMatch 442: flat-4	90.9%	9.1%	0.0%	39.1%
SmithsHeimann LS2: flat-4	94.3%	5.7%	0.0%	30.6%

(Table 5   BCII series, matches incorporating only one index finger, 4-finger)

[2-FINGER (2,3) vs IAFIS rolled]	TAR	1-TAR (FRR)	FAR	filter rate
BCII: rolled, mostly paper	97.3%	2.7%	1.9%	39.2%
Identix TP600/2000: rolled	97.8%	2.2%	1.7%	35.9%
CrossMatch 442: flat-4	90.9%	9.1%	1.2%	56.4%
SmithsHeimann LS2: flat-4	94.9%	5.1%	1.6%	48.5%

(Table 6   BCII series, matches incorporating only one index finger, 2-finger)

Performance penalties incurred when using plain-to-rolled identification

Tables 2 through 6 illustrate the other notable effect of the use of plain-to-rolled identification: a very significant increase in the filter rate.

The ATB, like most AFISes, removes from consideration a large fraction of the subjects in its database (the gallery) before embarking on the CPU-intensive task of feature-based matching. The ATB refers to the percentage **not** removed as the 'filter rate'. A filter rate of 2% means that only 24,000 subjects from a gallery of 1,200,000 were actually 'matched' against a particular probe. (We have found a correlation approaching 1 between the reported filter rate and the time taken to conduct a search. Filter rate in effect is the inverse of throughput; high filter rates mean slow performance.)

Viewing even the more optimistic (i.e., Smiths Heimann) case for plain fingerprints in Tables 2

and 3 shows, in both ten and eight finger matching, that a change from rolled to plain fingerprints carries with it very significant performance penalties of over 240% and almost 170%, respectively[n]. Table 4 shows that the equivalent penalty for two finger matching is 37%. Tables 5 and 6 show the penalties for a change from rolled to plain fingerprints for single-handed four finger matching to be 110%, and for single-handed two finger matching to be 39%.

Although the relative penalty for the shift from rolled to plain is smaller for reduced numbers of fingers, the absolute penalty is larger. In the case of reduced numbers of fingers, the baseline (i.e., rolled images) has already incurred a substantial penalty of its own: 200% for the reduction from ten to eight fingers, nearly 2200% for the reduction from ten to two fingers. This large absolute penalty is also noted for single-handed four finger and two finger matching.

Differences between the livescan devices

Tables 2 through 6 show that in every case the images acquired by the CrossMatch had a rate of failure to identify (1-TAR) that was between 60% and 100% higher than the rate for the images acquired by the Smiths Heimann. The same tables also show a filter rate between 15% and 60% higher for the CrossMatch. For these particular sets of images, the superiority for ATB matching of the set from the Smiths Heimann device is clear. The underlying reason is not clear.

The Smiths Heimann uses a larger platen. Matching becomes harder (slower) as the amount of data decreases, as shown by the degradation in throughput seen in tables 2 through 4 when the number of fingers is reduced and when the area of the images is reduced (from rolled to plain). Whether the difference in platen size causes the difference in accuracy and filter rate is uncertain.

Accuracy of the ATB as a model of IAFIS

Tables 7 and 8 compare the results of our ten and eight finger match experiments, using the ATB and its built-in background database of approximately 1.2M subjects, with the results of similar experiments conducted by the CJIS Division of the FBI, using the operational IAFIS with its live database of approximately 42M subjects.

That the filter rates are quite similar is evident; the rates from the ATB are in every case slightly smaller, averaging about 3.5% smaller than those from the IAFIS.

A similarity in failure rates is also evident; the rates from the ATB are in every case between 0% and 50% greater than those from the IAFIS, averaging about 15% larger (although the magnitude of the greatest difference is only 0.6%).

Thus the ATB with 1.2M subjects is a reasonable model of IAFIS with 42M subjects; the ATB's predictions of performance are a trifle optimistic, its predictions of accuracy a bit pessimistic.

---

[n] (e.g., (4.5 / 1.3) - 1 > 2.4)

Studies of Plain-to-Rolled Fingerprint Matching

[10-FINGER using ATB]	TAR	1-TAR	filter rate
Identix TP600/2000: rolled	97.9%	2.1%	1.3%
CrossMatch 442: flat-4	95.8%	4.2%	7.2%
SmithsHeimann LS2: flat-4	97.5%	2.5%	4.5%
[10-FINGER using IAFIS]			
IAFIS: Identix	98.3%	1.7%	1.4%
IAFIS: CrossMatch	95.9%	4.1%	7.5%
IAFIS: SmithsHeimann	98.1%	1.9%	4.6%

(Table 7   BCII series, comparing ATB to IAFIS, 10-finger)

[8-FINGER (2-5,7-10) using ATB]	TAR	1-TAR	filter rate
Identix TP600/2000: rolled	98.5%	1.5%	3.9%
CrossMatch 442: flat-4	96.5%	3.5%	15.5%
SmithsHeimann LS2: flat-4	98.2%	1.8%	10.5%
[8-FINGER using IAFIS]			
IAFIS: Identix	99.0%	1.0%	4.0%
IAFIS: CrossMatch	96.5%	3.5%	16.1%
IAFIS: SmithsHeimann	98.6%	1.4%	10.9%

(Table 8   BCII series, comparing ATB to IAFIS, 8-finger)

Greater inherent accuracy of rolled-to-rolled(?)

Consider the TARs shown in tables 2 through 4. The **m**ean TAR for **r**olled ten-finger prints, which we shall call TARm(r,10) is 97.5 [dropping the '%']; the **v**ariance TARv(r,10) is 0.8. The **m**ean TAR for **p**lain ten-finger prints, TARm(p,10) is 96.65; the **v**ariance TARv(p,10) is 1.7. We show these values for the ten-finger, eight-finger, and two-finger cases in table 9, where we also show the difference between the means (Var[iance] of Means) and the average value of the variances (Mean of Var[iance]s).

# fingers (n)	TARm(r,n)	TARv(r,n)	TARm(p,n)	TARv(p,n)	Var of Means	Mean of Vars
10	97.5	0.80	96.65	1.7	0.85	1.25
8	98.2	0.60	97.35	1.7	0.85	1.15
2	99.1	0.80	97.85	1.1	1.25	0.95

(Table 9   BCII series, comparing accuracy of rolled-to-rolled and plain-to-rolled)

We first note that in every case, given a number of fingers on which to match, the mean TAR for the rolled-to-rolled class of matching is higher than the mean TAR for the plain-to-rolled class of matching. That supports the hypothesis that the rolled-to-rolled class of matching is inherently more accurate, but fails to prove it. We also note that in two of the three instances there is less average variation between classes (Var of Means) than there is within class (Mean of Vars): the differences are greater than can be explained solely by a rolled-to-rolled versus plain-to-rolled classification. Even with extremely small classes (class-size = 2 for each class), the assertion that rolled-to-rolled is inherently more **accurate** is not proven.

Studies of Plain-to-Rolled Fingerprint Matching

5.2   Results from Group 2

Tables 10 through 15 below summarize the data obtained when matching the next three data sets, 2, 3, and 4. Two of the three types of matches shown (10- and 2-finger) are in nature directly comparable to BCII matches. The third type of match (6-finger) is in nature roughly comparable to the 8-finger BCII match.

[10-FINGER]	TAR	1-TAR (FRR)	FAR	filter rate
TXDPS: paper	90.4%	9.6%	0.0%	8.1%
INS10: livescan -> paper	82.1%	17.9%	0.0%	8.9%
INS_ben: livescan	89.1%	10.9%	0.0%	7.8%

(Table 10 TXDPS,INS10,INS_ben plain-to-rolled series, matches incorporating both index fingers, 10-finger)

[6-FINGER (2-4,7-9)]	TAR	1-TAR (FRR)	FAR	filter rate
TXDPS: paper	96.3%	3.7%	0.1%	12.9%
INS10: livescan -> paper	86.9%	13.1%	0.1%	15.2%
INS_ben: livescan	87.6%	12.4%	0.0%	17.8%

(Table 11   TXDPS,INS10,INS_ben plain-to-rolled series, matches incorporating both index fingers, 6-finger)

[6-FINGER (1-3,6-8)]	TAR	1-TAR (FRR)	FAR	filter rate
TXDPS: paper	92.1%	7.9%	0.0%	17.5%
INS10: livescan -> paper	84.5%	15.5%	0.1%	18.6%
INS_ben: livescan	90.5%	9.5%	0.0%	14.2%

(Table 12 TXDPS,INS10,INS_ben plain-to-rolled series, matches incorporating both index fingers, 6-finger)

[4-FINGER (2-3,7-8)]	TAR	1-TAR (FRR)	FAR	filter rate
TXDPS: paper	97.0%	3.0%	1.4%	21.8%
INS10: livescan -> paper	88.2%	11.8%	1.2%	24.7%
INS_ben: livescan	89.4%	10.6%	0.6%	26.8%

(Table 13 TXDPS,INS10,INS_ben plain-to-rolled series, matches incorporating both index fingers, 4-finger)

[4-FINGER (1-2,6-7)]	TAR	1-TAR (FRR)	FAR	filter rate
TXDPS: paper	92.0%	8.0%	0.4%	30.9%
INS10: livescan -> paper	83.7%	16.3%	0.4%	32.8%
INS_ben: livescan	90.5%	9.5%	0.3%	23.3%

(Table 14 TXDPS,INS10,INS_ben plain-to-rolled series, matches incorporating both index fingers, 4-finger)

[2-FINGER (2,7)]	TAR	1-TAR (FRR)	FAR	filter rate
TXDPS: paper	97.4%	2.6%	44.7%	38.1%
INS10: livescan -> paper	88.1%	11.9%	35.3%	42.8%
INS_ben: livescan	89.7%	10.3%	29.2%	44.4%

(Table 15 TXDPS,INS10,INS_ben plain-to-rolled series, matches incorporating both index fingers, 2-finger)

The data sets forming group 2, sets 2, 3, and 4, are not homogeneous. The inhomogeneity is more evident in accuracy (TAR) than it is in filter rate.

Although the TXDPS data exhibits an accuracy not far removed from that of the plain subsets of group 1 (i.e., BCII subsets 1.3 and 1.4), both the INS10 and the INS_ben data exhibit markedly lower accuracy. This is particularly evident in the 6-finger and 2-finger results.

However, it should be noted that the filter rates of sets 2, 3, and 4 are quite similar for 10-finger, 6-finger, and 2-finger results. It should also be noted that both the 6-finger and 2-finger filter rates are comparable to the 8-finger and 2-finger rates for the plain subsets of group 1.

These dissimilarities aside, it should be noted that of the 18 available comparisons[o] of TAR($\mathbf{s}$,n) with TAR($\mathbf{s}$,$m<n$), where $\mathbf{s}$ is set and n is finger count, in 16 cases TAR($\mathbf{s}$,n) < TAR($\mathbf{s}$,m). This is the same phenomenon that was noted within group 1: TAR decreases as the number of fingers increases.

The 4-finger statistics are, as one might expect, intermediate between the 2-finger and 6-finger statistics. Although the group 2 statistics are not strictly equivalent to the 4-finger statistics from group 1 (the 4-finger group 1 results were based on four fingers from a single hand), the values roughly agree. Of interest is the comparison between tables 11 and 12, and also between 13 and 14. In both cases the use of the thumb (tables 12 and 14) produced a modest improvement for INS_ben data, but an even greater degradation for both INS10 and TXDPS data.

Not all matchers had such mediocre results with sets 3 and 4 [4]. The qualities which cause data to match well may vary from matcher to matcher.

Effect of gallery size on matcher accuracy

Other results from unpublished earlier studies involving the sets in group 2 support our conclusion that TAR is largely independent of the size of the (seeded) gallery while supporting the conclusion that FAR is a positive monotonic function of (unseeded) gallery size. These are summarized below in tables 16 through 18.

gallery size (seeded)	filter rate	seconds per probe	:(TAR)	Φ(TAR)	:(FAR)	Φ(FAR)
1229 K	8.1%	2.03	90.4%	0.34	0.035%	0.020
640 K	8.1%	1.13	90.5%	0.32	0.015%	0.020
320 K	8.5%	0.53	90.6%	0.34	0.0050%	0.0081
160 K	8.3%	0.27	90.6%	0.33	0.0017%	0.0053
80 K	8.6%	0.17	90.6%	0.29	0.0017%	0.0053

(Table 16   effects of gallery size on matching accuracy and performance: TXDPS data)

---

[o] viewing the most conventional cases: 10-finger; 6-finger (2-4,7-9); 4-finger (2-3,7-8); and 2-finger (2-7)

Studies of Plain-to-Rolled Fingerprint Matching

gallery size (seeded)	filter rate	seconds per probe	:(TAR)	Φ(TAR)	:(FAR)	Φ(FAR)
1223 K	8.9%	2.43	82.1%	0.69	0.042%	0.024
634 K	8.8%	1.26	82.2%	0.67	0.025%	0.024
314 K	9.1%	0.62	82.4%	0.64	0.015%	0.014
154 K	9.2%	0.32	82.5%	0.61	0.0083%	0.0089
74 K	9.8%	0.19	82.6%	0.66	0.0062%	0.0086

(Table 17   effects of gallery size on matching accuracy and performance: INS10 data)

gallery size (seeded)	filter rate	seconds per probe	:(TAR)	Φ(TAR)	:(FAR)	Φ(FAR)
1229 K	7.8%	1.93	89.1%	0.60	0.0050%	0.0081
640 K	7.7%	0.94	89.4%	0.28	0.0034%	0.0071
320 K	7.8%	0.49	89.6%	0.30	0.0017%	0.0053
160 K	8.0%	0.26	89.8%	0.28	0.0000%	0.0000
80 K	8.4%	0.16	90.0%	0.31	0.0000%	0.0000

(Table 18   effects of gallery size on matching accuracy and performance: INS_ben data)

When the total gallery size is 80K (74K), all but 20K comes from the seeded mates of the probe set; at this point, the subject data set rather than the IAFIS Criminal Master File (CMF) predominates in the gallery. We can thus ignore anomalies introduced by galleries of this size.

Figure 2 shows that the time required to process each probe is a linear function of the 'effective gallery size', which is the product of the gallery size and the filter rate.

Figure 2 - search time as a function of effective gallery size
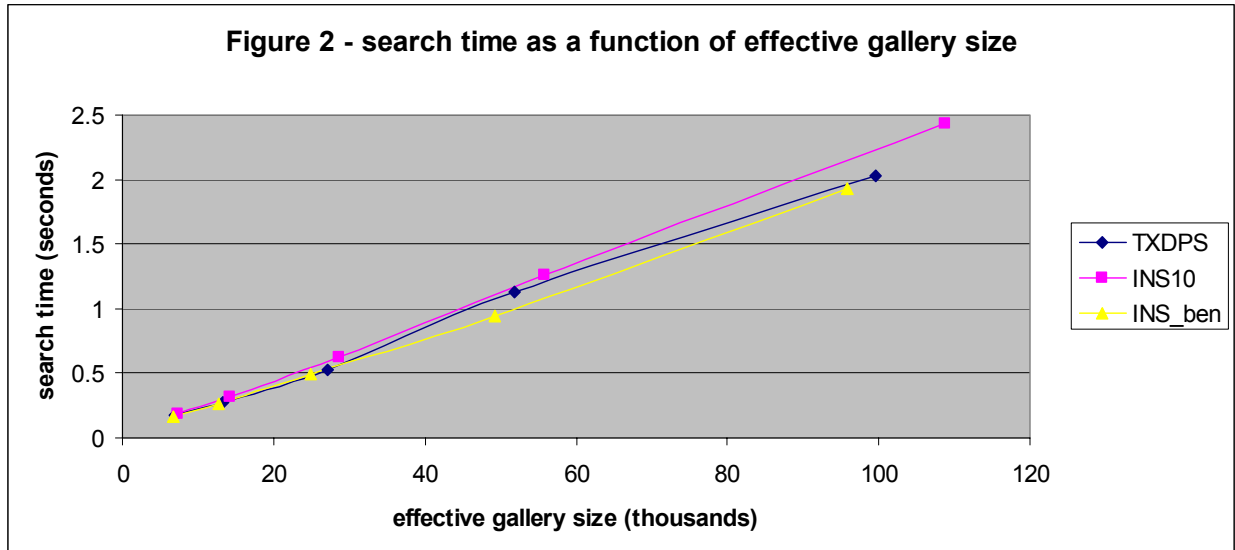
Studies of Plain-to-Rolled Fingerprint Matching

Figure 3 shows that the rate of successful identification, TAR, is essentially independent of gallery size.

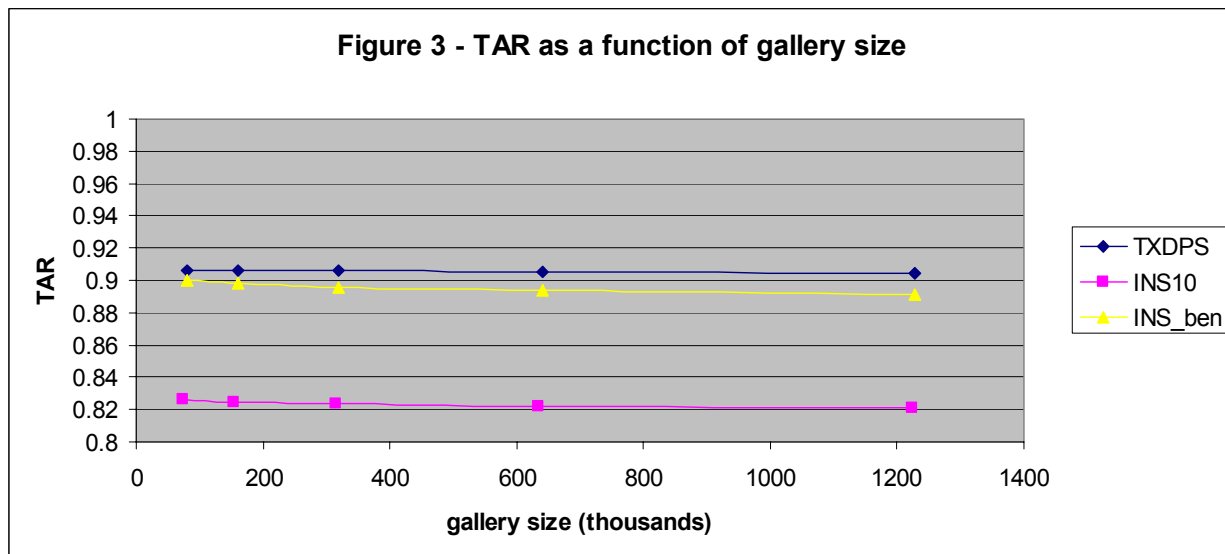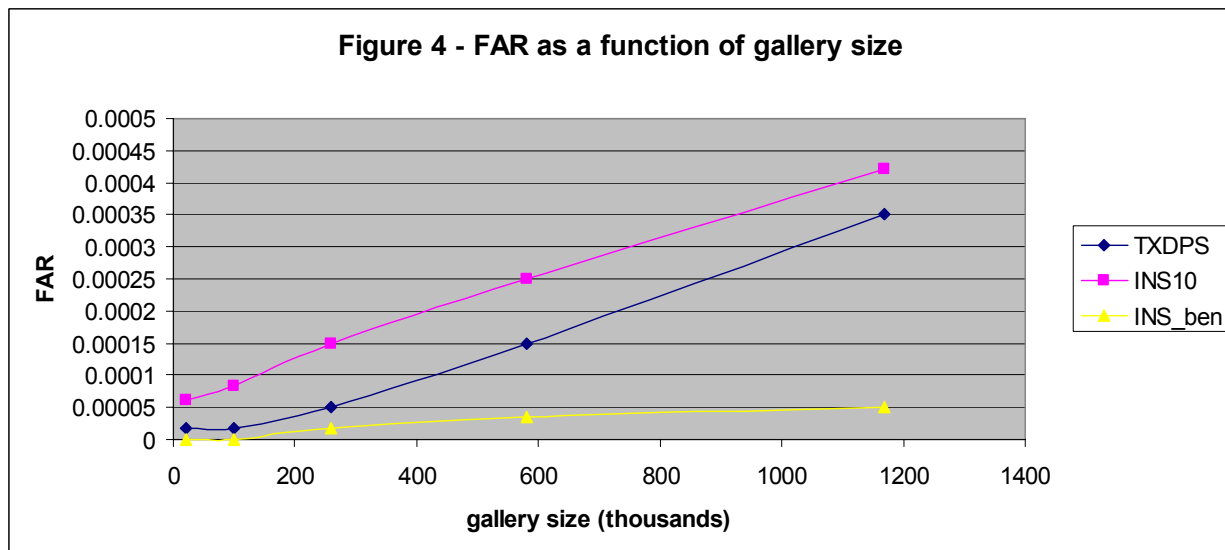Figure 3 - TAR as a function of gallery size

Figure 4 shows that the rate of mistaking a subject not in the gallery for a subject in the gallery, FAR, increases as the size of the unseeded composite gallery increases.

Figure 4 - FAR as a function of gallery size

5.3 Results from Group 3

Group 3 includes the Ohio BCII plain subsets and the large INS_plain and BCC sets

5.3.1 Results from group 3 - BCII plain subsets

We enrolled each of the five **plain** fingerprint subsets of set 1 (Ohio BCII data) in the empty repository, resulting in an unconsolidated target with 4175 records but only 925 subjects. This target physically was one unconsolidated gallery. We treated it logically as five disjoint, fully consolidated galleries.

We used each plain subset as a probe to search against each of the other subsets as a gallery. To accomplish this, we searched with each probe against the entire target, ignoring all match scores that did not correspond to the desired gallery set. We conducted those searches in five different modes: 10-finger, 8-finger, 2-finger (both index fingers), 4-finger (right hand only), and 2-finger (right index and right middle). Only those subjects common to a probe-gallery combination were considered in scoring and analyzing that probe-gallery combination.

Tables 19 through 23 show the results of the searches described above and in 4.3 (qv). The five different search modes (10-, 8-, 2-, 4-, and 2-finger, as described above), are shown in the bottom left entry of each table.

The mean values for FAR and TAR were computed by **probe** subset, defined as the mean FAR (TAR) over all galleries other than the gallery that also constitutes the probe:
$$FAR_{Pi} = 3\ FAR(P_i, G_{k\ i}) / 4 \text{ and } TAR_{Pi} = 3\ TAR(P_i, G_{k\ i}) / 4$$

Likewise, FAR and TAR means were computed by **gallery** subset:
$$FAR_{Gi} = 3\ FAR(G_i, P_{k\ i}) / 4 \text{ and } TAR_{Gi} = 3\ TAR(G_i, P_{k\ i}) / 4$$

For each subset *i*, the mean of its two FARs and TARs were taken, and are shown in the line labelled 'probe,gal mean': $(FAR_{Pi} + FAR_{Gi}) / 2$ and $(TAR_{Pi} + TAR_{Gi}) / 2$. These are accuracy metrics for that subset (e.g., CrossMatch) for that class of matching (e.g., ten-finger).

The means of all such FAR and TAR are shown in the last line. They are overall accuracy metrics for that class of matching. The aggregate filter rate is also shown in the last line of each table. It is an inverted performance (throughput) metric for that class of matching: the higher the value, the lower the throughput.

The plain-to-plain searches were conducted without the 1.2M IAFIS background, which consists of rolled images. Although the TAR values thus obtained may fairly be compared to the TAR values of the equivalent plain-to-rolled searches, there is no assurance that the FAR values are also comparable: the gallery sizes differ by three orders of magnitude.

Studies of Plain-to-Rolled Fingerprint Matching

Table 19 shows that in the aggregate, plain-to-plain ten-finger identification matching using the ATB was substantially less accurate than plain-to-rolled matching (see table 2). The most interesting value in table 19 is the mean TAR achieved, 90.3%, using the IAFIS master plain subset as the gallery (because the IAFIS master rolled subset was used to seed the gallery in the plain-to-rolled matching, we used the corresponding plain as the basis of comparison). The resultant value for (1-TAR), 9.7%, is 2.9 times the mean, 3.35%, of the two plain-to-rolled (1-TAR) values in table 2.

Table 19 also shows that plain-to-plain matching took 50% more resources than plain-to-rolled matching. Table 2 shows a mean filter rate for plain-to-rolled matching of less than 6%, whereas Table 19 shows a plain-to-plain filter rate of 9%.

However, when we look at the same data another way, we find that when we use only plain fingerprint records that were originally built as plain fingerprint records (i.e., only CrossMatch and Smiths Heimann records), we achieve a value for (1-TAR) of 1.7%. This represents a reduction in error by more than half an order of magnitude, and is **better** than **any** of the plain-to-rolled ten-finger results (cf table 2). [The results on the main diagonal (e.g., Identix:Identix) are almost perfect. This has no real significance: it would be notable only if it were not true.]

At this point we simply haven't enough data to explain this disparity. Although we can suggest some likely causes, that would provide speculation rather than information. However, both the magnitude of the disparity and its potential operational implications warrant further study.

GALLERY:	BCII master		Identix		CrossMatch		SmithsHeimann		IAFIS master	
	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
PROBE:										
BCII master	100.0%	0.0%	88.8%	0.0%	89.8%	0.0%	91.1%	0.0%	87.7%	0.0%
Identix	88.3%	0.0%	100.0%	0.0%	89.9%	0.0%	91.2%	0.0%	88.5%	0.0%
CrossMatch	92.5%	0.0%	92.5%	0.0%	100.0%	0.0%	98.3%	0.0%	91.5%	0.0%
SmithsHeimann	93.8%	0.0%	93.9%	0.0%	98.3%	0.0%	99.9%	0.0%	93.3%	0.0%
IAFIS master	88.6%	0.0%	89.9%	0.0%	89.4%	0.0%	91.4%	0.0%	100.0%	0.0%
gallery mean	90.8%	0.0%	91.3%	0.0%	91.8%	0.0%	93.0%	0.0%	90.3%	0.0%
probe mean	89.4%	0.0%	89.5%	0.0%	93.7%	0.0%	94.8%	0.0%	89.8%	0.0%
probe,gal mean	90.1%	0.0%	90.4%	0.0%	92.8%	0.0%	93.9%	0.0%	90.0%	0.0%
fingers 1-5,6-10	TAR	91.4%			FAR	0.0%			filter	9%

(Table 19   Ohio BCII series, plain-to-plain matches incorporating both index fingers, 10-finger)

Another statistic of interest is the mean TAR for each livescan probe set (i.e., Identix, CrossMatch, Smiths Heimann) for the card-based gallery sets (i.e., BCII master, IAFIS master)

compared to the mean TAR of that same probe set for the livescan gallery sets. The former means are, respectively, 88.4%, 92.0%, and 93.6%, whereas the latter are 90.6%, 95.4%, and 96.1%, an absolute improvement of 2.7%. (The results are more dramatic when we reverse the role of probe and gallery; in this case the former means are, respectively, 89.4%, 89.6%, and 91.2%, whereas the latter are 93.2%, 94.1%, and 94.8%, an absolute improvement of nearly 4%.)

Accuracy of plain-to-rolled identification with fewer than ten fingers

Tables 20 and 21 show that plain-to-plain identification matching using fewer than 10 fingers can also be quite accurate.

Note that the overall TAR shown in table 20, using eight fingers, was 93.9%. At first glance, this is much inferior to the results shown in table 3. Just as with the ten-finger results shown in table 19, when we use only plain fingerprint records that were originally built as plain fingerprint records, we achieve a value for (1-TAR) of 1.3%. This is superior to any of the plain-to-rolled ten-finger results. Here we are also struck by a marked asymmetry: using the Identix images in the probe and either the CrossMatch or Smiths Heimann in the gallery, the (1-TAR) figures were not particularly good; the mean was 5.9%. However, when the roles were reversed, there was a dramatic improvement; the mean decreased to 1.8%. Again, further investigation is needed.

When we consider the two-finger results shown in table 21, we no longer see any asymmetry in the Identix-to-CrossMatch or Identix-to-Smiths Heimann data. We do see that the best results are superior to the best two-finger plain-to-rolled results (cf table 4), although in this case the best TAR values, 99.7%, are provided by pairing Identix and Smiths Heimann images. As with table 20, however, the overall TAR, 97%, is inferior to any of the two-finger plain-to-rolled results (cf table 4). Although the FAR values appear to be remarkably improved, we have already noted that the change in gallery size makes this apparent improvement signify little.

When we repeat for eight fingers the comparison of the mean TAR for each livescan probe set for the card-based gallery set (i.e., BCII master, IAFIS master) to the mean TAR for the livescan gallery set, we find that means of 91.2%, 94.7%, and 96.4% have increased to 94.1%, 98%, and 98.8%, an improvement of almost 3%. (Again, the results are more dramatic when we reverse the role of probe and gallery; in this case the former means are, respectively, 92.2%, 90.9%, and 93.1%, whereas the latter are 98.2%, 96.2%, and 96.6%, an absolute improvement of nearly 5%.)

In the corresponding case for two fingers, the improvement is even more marked: from 96.8%, 95.1%, and 96.8% to 99.2%, 98.9%, and 99.4%, reducing the failure rate from 3.8% to 0.8%. (Once again, the results are also dramatic when we reverse the role of probe and gallery; in this case the former means are, respectively, 97.1%, 95.2%, and 96.8%, whereas the latter are 99.2%, 99.0%, and 99.4%, reducing the failure rate from 3.6% to 0.7%.)

Studies of Plain-to-Rolled Fingerprint Matching

GALLERY:	BCII master		Identix		CrossMatch		SmithsHeimann		IAFIS master	
	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
PROBE:										
BCII master	99.5%	0.0%	90.4%	0.0%	90.0%	0.0%	91.6%	0.0%	89.0%	0.0%
Identix	90.4%	0.0%	99.8%	0.0%	93.6%	0.0%	94.6%	0.0%	92.1%	0.0%
CrossMatch	94.3%	0.0%	97.5%	0.0%	99.9%	0.0%	98.6%	0.0%	95.1%	0.0%
SmithsHeimann	95.7%	0.0%	98.9%	0.0%	98.8%	0.0%	99.7%	0.0%	97.2%	0.0%
IAFIS master	90.1%	0.0%	93.9%	0.0%	91.8%	0.0%	94.6%	0.0%	99.7%	0.0%
gallery mean	92.6%	0.0%	95.2%	0.0%	93.5%	0.0%	94.8%	0.0%	93.3%	0.0%
probe mean	90.3%	0.0%	92.7%	0.0%	96.3%	0.0%	97.7%	0.0%	92.6%	0.0%
probe,gal mean	91.4%	0.0%	93.9%	0.0%	94.9%	0.0%	96.2%	0.0%	93.0%	0.0%
fingers 2-5,7-10	TAR	93.9%			FAR	0.0%			filter	15%

(Table 20   Ohio BCII series, matches plain-to-plain incorporating both index fingers, 8-finger)

GALLERY:	BCII master		Identix		CrossMatch		SmithsHeimann		IAFIS master	
	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
PROBE:										
BCII master	100.0%	0.0%	96.8%	0.2%	95.2%	0.0%	96.2%	0.3%	94.5%	0.0%
Identix	96.6%	0.0%	99.9%	0.1%	98.8%	0.0%	99.7%	0.1%	97.0%	0.3%
CrossMatch	95.1%	0.0%	98.8%	0.0%	99.8%	0.1%	99.0%	0.1%	95.1%	0.0%
SmithsHeimann	96.3%	0.0%	99.7%	0.1%	99.2%	0.2%	99.8%	0.2%	97.4%	0.3%
IAFIS master	94.3%	0.0%	97.4%	0.1%	95.1%	0.1%	97.5%	0.3%	99.9%	0.1%
gallery mean	95.6%	0.0%	98.2%	0.1%	97.1%	0.1%	98.1%	0.2%	96.0%	0.1%
probe mean	95.7%	0.1%	98.0%	0.1%	97.0%	0.0%	98.2%	0.1%	96.1%	0.1%
probe,gal mean	95.6%	0.1%	98.1%	0.1%	97.0%	0.1%	98.1%	0.2%	96.0%	0.1%
fingers 2,7	TAR	97.0%			FAR	0.1%			filter	46%

(Table 21   Ohio BCII series, plain-to-plain matches incorporating both index fingers, 2-finger)

As tables 22 and 23 demonstrate, failure to use both index fingers results in a significant drop in accuracy, a pattern previously evident in tables 5 and 6.

GALLERY:	BCII master		Identix		CrossMatch		SmithsHeimann		IAFIS master	
	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
PROBE:										
BCII master	99.8%	0.0%	91.1%	0.0%	86.6%	0.0%	91.3%	0.0%	84.0%	0.0%
Identix	90.4%	0.0%	100.0%	0.0%	94.0%	0.0%	97.8%	0.0%	87.3%	0.0%
CrossMatch	86.8%	0.0%	95.5%	0.0%	99.6%	0.0%	95.8%	0.0%	84.3%	0.0%
SmithsHeimann	91.0%	0.0%	98.7%	0.0%	95.2%	0.0%	100.0%	0.0%	89.5%	0.0%
IAFIS master	84.2%	0.0%	88.9%	0.0%	85.0%	0.0%	88.5%	0.0%	99.9%	0.0%
gallery mean	88.1%	0.0%	93.5%	0.0%	90.2%	0.0%	93.3%	0.0%	86.3%	0.0%
probe mean	88.3%	0.0%	92.4%	0.0%	90.6%	0.0%	93.6%	0.0%	86.6%	0.0%
probe,gal mean	88.2%	0.0%	93.0%	0.0%	90.4%	0.0%	93.5%	0.0%	86.5%	0.0%
fingers 2,3,4,5	TAR	90.3%			FAR	0.0%			Filter	38%

(Table 22　　Ohio BCII series, plain-to-plain matches incorporating one index finger, 4-finger)

GALLERY:	BCII master		Identix		CrossMatch		SmithsHeimann		IAFIS master	
	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR	TAR	FAR
PROBE:										
BCII master	100.0%	0.0%	92.5%	0.3%	87.8%	0.0%	91.8%	0.2%	85.7%	0.2%
Identix	92.1%	0.0%	100.0%	0.0%	95.0%	0.0%	98.9%	0.1%	88.7%	0.1%
CrossMatch	87.0%	0.0%	95.1%	0.0%	99.6%	0.0%	96.1%	0.0%	84.7%	0.0%
SmithsHeimann	90.6%	0.2%	98.5%	0.2%	95.2%	0.1%	99.7%	0.3%	89.3%	0.1%
IAFIS master	85.0%	0.4%	90.0%	0.0%	85.7%	0.1%	90.3%	0.1%	99.9%	0.1%
gallery mean	88.7%	0.1%	94.0%	0.1%	90.9%	0.1%	94.3%	0.1%	87.1%	0.1%
probe mean	89.4%	0.2%	93.7%	0.1%	90.7%	0.0%	93.4%	0.2%	87.7%	0.2%
probe,gal mean	89.0%	0.1%	93.9%	0.1%	90.8%	0.0%	93.8%	0.1%	87.4%	0.1%
fingers 2,3	TAR	91.0%			FAR	0.1%			Filter	54%

(Table 23　　Ohio BCII series, plain-to-plain matches incorporating one index finger, 2-finger)

Performance penalties incurred when using plain-to-plain identification

Tables 19 through 21 (above) illustrate that the use of plain-to-plain matching for identification causes, just as did the use of plain-to-rolled matching, a very significant increase in the filter rate. This can most readily be seen by examining the filter rates for these tables: 9%, 15%, and 46%, which are marginally higher than the mean plain-to-rolled filter rates of 6%, 13%, and 46% shown in tables 2 through 4.

5.3.2 Results from group 3 - large sets of two-finger plain

Tables 24 and 25 below summarize the data obtained when matching the last two data sets, 5 and 6. The results can most instructively be compared to those in table 21. It is also useful to compare them to the plain-to-rolled results in table 15.

The filter rates of 45% to 49% agree quite closely with the 46% obtained in table 21, and are not far removed from the 38% to 44% reported in table 15. The TARs, around 94% in both tables 24 and 25, are somewhat inferior to all of those in table 21, and bracketed by the 88% to 97% reported in table 15.

There is rough agreement; that it is rough should not surprise us, given the disparate nature of the data sources.

Note that these two tables are instructive in themselves. First we observe that the combination of the INS_plain and BCC data in the 574K gallery introduced no anomalies; the results are reasonable, and what one might expect from increasing the size of either gallery with more images from its own source. Next we observe that, as expected, FAR does increase with gallery size. Third, we note in table 25 that use of a non-random probe set can skew the results: compare the FAR of 4.3% in line 1 (60K probe selected at **random** from all 274K subjects) to the FAR of 5.8% in line 2 (60K probe consisting of the **first** 60K subjects).

probe / gallery	filter rate	seconds per probe	:(TAR)	Φ(TAR)	:(FAR)	Φ(FAR)
1st 60K / 60K INS_plain	48.7%	<1	94.4%	0.34	0.74%	0.094
1st 60K / 574K combined	48.1%	4.04	94.0%	0.24	9.88%	0.424

(Table 24   effects of gallery size & composition on matching accuracy and performance: INS_plain data)

probe / gallery	filter rate	seconds per probe	:(TAR)	Φ(TAR)	:(FAR)	Φ(FAR)
random 60K / 274K BCC	44.8%	1.64	94.0%	0.21	4.3%	0.23
1st 60K / 274K BCC	44.8%	1.64	93.9%	0.27	5.8%	0.31
1st 60K / 574K combined	45.8%	3.79	93.9%	0.28	8.0%	0.34

(Table 25   effects of gallery size & composition on matching accuracy and performance: BCC data)

5.4   The Choice of Fingers

The ATB offers important choices. Whether to match using plain or rolled images is one. How many fingers to use is another. The selection of which fingers to use is a third.

Studies of Plain-to-Rolled Fingerprint Matching

The ATB, like IAFIS, has a strong preference for index fingers.

Tables 5 and 6 illustrate this, the former very dramatically. Table 5, with four fingers (2,3,4,5) available, shows a failure rate (1-TAR) between 2.4 and 3.5 times as great as table 4, with only two fingers (2,7) available; however, it has a noticeably better filter rate. Table 6, with the same number of fingers as table 4 but with only one index finger (2,3), shows failure rates similar to table 5, with filter rates only moderately poorer than table 4.

This preference is also evident in plain-to-plain matching. Tables 22 and 23, when compared to tables 20 and 21, show a marked drop in accuracy when only one index finger is used. This most evident when comparing 21 and 23: the overall failure rate (1-TAR) for two-finger matches is 3% in the former case, 9% in the latter.

This supports the view of the ATB (and IAFIS) as an index finger matcher preceded by a ten finger filter.

A matter for investigation is whether this preference for index fingers is a property of fingers or of the ATB: how well would the ATB perform if it were presented with finger 2 images in the finger 2 position but with finger 3 images in the finger 7 position (i.e., if we were to mislead it)?

6.   INTERPRETATION / SUMMARY / CONCLUSIONS

We asked nine questions. We have answered all of them:

(1) how accurate is plain-to-rolled identification? is it less accurate than rolled-to-rolled?
Failure to identify (1-TAR) can be 2% or less. Accuracy can be as good as with rolled prints of average quality.

(2) how accurate is plain-to-rolled identification with fewer than ten fingers?
Accuracy with fewer than ten fingers can exceed accuracy with ten fingers, but performance penalties can be expected (see next question).

(3) are performance penalties incurred by adopting plain-to-rolled identification? if penalties are incurred, how great are they?
Performance (throughput) deteriorates markedly as the amount of data is reduced. Using plain rather than rolled fingerprints in 10-finger or 8-finger matching can be expected to result in a three-fold reduction in throughput. Reducing the number of fingers from ten to two can be expected to reduce throughput by a factor of ten or more. These effects are independent of one another, and thus multiplicative.

(4) are there significant differences in accuracy and performance between the fingerprint images supplied by two different models of livescan devices? if there are differences, how great are they?
There are significant differences in both matching accuracy and performance. In this case, the inferior set of images showed a rate of failure to identify (1-TAR) 68% to 94% higher than that shown by the superior, and a filter rate 24% to 60% higher.

(5) does gallery size affect accuracy? if there is an effect, how can it be characterized?
Within the range of gallery sizes covered by these studies (160,000 to 42,000,000), gallery size did not affect overall accuracy (TAR).

(6) does the ATB accurately model the behavior of IAFIS?
In a word, yes.

(7) how accurate is plain-to-plain identification?
Failure to identify (1-TAR) can be 2% or less. There is more variability than evident with rolled-to-plain.

(8) how accurate is plain-to-plain identification with fewer than ten fingers?
As with rolled-to-plain, accuracy with fewer than ten fingers can exceed accuracy with ten fingers, with a failure rate of well under 1%. There is more variability than evident with rolled-to-plain, and performance penalties can be expected.

Studies of Plain-to-Rolled Fingerprint Matching

(9) are performance penalties incurred by adopting plain-to-plain identification?
Yes. Filter rates can approach 50%, which means matching every other subject. (Compared to the 2% filter rate typical of rolled-to-rolled tenprint matching, a 50% rate implies a 25-fold increase in matching workload.)

We were able to draw other conclusions as well:

There were clear differences, but the interclass difference (between rolled and plain) was smaller than the intraclass differences (between BCII and Identix and between CrossMatch and Smiths Heimann).

There were similarities among all the sets; the invariance of TAR over gallery size and the essentially linear relationship of FAR to gallery size were both notable.

The ATB works most accurately (i.e., has the lowest failure rate, or (1-TAR)) with two index fingers. The addition of more fingers raises its throughput markedly but with a concomitant reduction in TAR. The use of six or more fingers reduces the FAR to almost zero.

7.   FUTURE INVESTIGATION

The major topics to be covered in this study have been addressed.

There are details that can be added.

*   We intend to develop best-fit and best-fit-linear equations for TAR and filter-rate
*   We intend to investigate behavior when 2,3 is treated as if it were 2,7

However, the nature of the qualities in fingerprint images needed for consistently successful matching with two plain index fingers **requires** further investigation.

REFERENCES

 (1)   Public Law 107-56 (USA PATRIOT ACT); 107th United States Congress, Washington, D.C.; 26 October 2001

(2)   Public Law 107-173 (Enhanced Border Security and Visa Entry Reform Act of 2002); 107th United States Congress, Washington, D.C.; 14 May 2002

(3)   M D Garris, C I Watson, R M McCabe, C L Wilson; "User's Guide to NIST Fingerprint Image Software (NFIS)" NISTIR 6813; National Institute of Standards and Technology, Gaithersburg, Maryland; August 2001

(4)   C L Wilson, C I Watson, M D Garris, A Hicklin; "Studies of Fingerprint Matching Using the NIST Verification Test Bed (VTB)" NISTIR 7020; National Institute of Standards and Technology, Gaithersburg, Maryland; 07 July 2003

 (5)   Federal Bureau of Investigation; "The Science of Fingerprints: Classification and Uses"; U.S. Government Printing Office, Washington D.C.; 1984.

 (6)   "NIST Algorithm Test Bed (ATB) User Manual"; Lockheed Martin Corporation, Orlando, Florida; 12 September 2002

 (7)   Daubert v. Merrell Dow Pharmaceuticals (92-102); United States Supreme Court, Washington, D.C.; 28 June 1993; <http://supct/law/cornell.edu/supct/html/92-102.ZS.html> (as of 20040123)

(8)   Kumho Tire Co. v. Carmichael (97-1709); United States Supreme Court, Washington, D.C.; 23 March 1999; <http://supct/law/cornell.edu/supct/html/97-1709.ZS.html> (as of 20040123)

(9)   J L Wayman; "When Bad Science Leads to Good Law: the Disturbing Irony of the *Daubert* Hearing in the Case of *U.S. v Byron C. Mitchell*"; San Jose State University, San Jose, California; 02 February 2000; <http://www.engr.sjsu.edu/biometrics/publications_daubert.html> (as of 20040123)

(10)   S A Rizvi, P J Phillips, H Moon; "The FERET Verification Testing Protocol for Face Recognition Algorithms" NISTIR 6281; National Institute of Standards and Technology, Gaithersburg, Maryland; October 1998

(11)   J D Murray, W vanRyper; "Encyclopedia of Graphics File Formats"; O'Reilly and Associates, Inc., Sebastopol, California; 1994

(12)   "Electronic Fingerprint Transmission Specification" CJIS-RS-0010 (V7); Criminal

Justice Information Services Division, Federal Bureau of Investigation, Washington, D.C.; 29 January 1999

(13)   R M McCabe; "Data Format for the Interchange of Fingerprint, Facial, Scar Mark & Tattoo (SMT) Information"; American National Standard ANSI/NIST-ITL 1-2000; National Institute of Standards and Technology , Gaithersburg, MD; July 2000

(14)   "National WebCheck Program Pilot Project Preliminary Report"; Ohio Attorney General Bureau of Criminal Identification and Investigation; London, Ohio; December 2003

Studies of Plain-to-Rolled Fingerprint Matching

APPENDIX A:

ATB SETTINGS

The settable search characteristics used were all default values:

TP CMF Caxi Filter Rate:	0.01
TP CMF ARG Low Threshold:	2000
TP CMF ARG High Threshold:	16000
TP CMF ARG Final Threshold:	3000
TP CMF Filtering Threshold:	1.1216

The settable quality characteristics used were all default values except as noted:

Missing Finger Threshold:	9	
Criminal TP Search Reject Threshold:	81	[from LM; default was 69]

In addition, LM stated that it had made the following modifications (to improve usability in the NIST environment) before delivering the ATB, none affecting match scores and only one potentially affecting performance (throughput):

- Disabled FNU validation – this allows subjects to be seeded into the repository without having to meet the normal FNU naming conventions

- Disabled loading of Civil, SLC and ULF repositories – this allows the system to run the AFIS Segment[p] Control and Processing Server functions on a single server (reduces the amount of memory consumed)

- Increased the number of TP search candidates passed back to Segment Control to 30 – by setting the tunable ARG Final threshold to 0, this enables NIST to receive the top 30 TP candidates versus just the candidates the AFIS segment would normally nominate.

- Disabled automatic cascading of 1 to 1 Verification to TP search – normally if a 1 to 1 Verification submission does not result in the provided FNU matching the provided imagery, a normal TP search is automatically performed. This was disabled so that just the 1 to 1 Verification could be analyzed.

- Added the capability to disable RRI filtering – an environment variable was added that will allow the AFIS segment to perform TP/TP searches without the RRI filter. The normal AFIS segment does not provide a mechanism to disable this filter.

---

[p] in Appendix A, a segment is a portion of IAFIS; the three major segments are known as AFIS, III, and ITN

Studies of Plain-to-Rolled Fingerprint Matching

APPENDIX B:

POSSIBLE BIASES: DATA SET 1 (BCII)

(A) Only those subjects (1) who were incarcerated, and (2) for whom useable fingerprint images were on file, were used to provide data for this study. Although it is not known whether condition (2) was a selection criterion, condition (1) was necessarily a criterion. If one makes the reasonable assumption that criminals who have 'good quality' (for our purposes, this means matchable using existing AFIS technology) fingerprints are more likely to be apprehended and thus incarcerated than criminals who have 'poor quality' fingerprints, then there is some bias in this study: it is a measure of the most effective way to keep on apprehending those whom we now apprehend. This is not necessarily an unwanted result.

The reasonableness of the assumption can be tested by denying it; to deny it would be to deny that the operation of IAFIS aids in the apprehension and incarceration of criminals.

(B) Because the BCII data were collected for a special study, the data may not be truly representative of operational data with all its warts and blemishes.

(C) The criteria used for selecting the subjects from the available incarcerated prisoners are not known to the NIST investigators.

(D) If TAR is dependent on the size of the gallery $c(G)$, then the results might not extrapolate well from the ATB to the gallery(ies) of the actual application(s). However, tables 6 and 7 showed that any such effects **in this case** would be small.

(E) There was no evident randomization of the order in which a given subject was fingerprinted on a given type of livescan machine (Identix, CrossMatch, Smiths Heimann). It is not inconceivable that subjects tended to become {more | less} cooperative with each subsequent fingerprinting.

(F) The world of operational fingerprint identification is largely concerned with matching a subject's recent images with those acquired at some time in the past. Should there be a short-term (short-lived) component of matchable features (e.g., caused by nicks and scrapes), that component might add a bias favoring images acquired at nearly the same time.

From a cursory inspection of the 'Date Printed' [2.038] and 'Date' [1.05] fields in each record, the Smiths Heimann samples were obtained on July 1, 2003, whereas both the Identix and CrossMatch samples were obtained on June 25-30, 2003. However, it is also evident that the data in the records is not pristine, so it is reasonable that the dates may not accurately reflect the date when the subject was printed. (The protocol used to create and process the samples has **not** been requested from the Ohio BCII.)

Studies of Plain-to-Rolled Fingerprint Matching

POSSIBLE BIASES: DATA SETS 2, 3, 4, 5, 6

(A) The data sets used excluded records for which automatic segmentation failed to extract eight plain fingerprint images. This in turn may have tended to exclude subjects whose fingerprint images are not distinct.

(B) The automatic classification of potential consolidations as confirmed consolidations (and their subsequent removal) may have reduced the frequency of false matches.

APPENDIX C:

WHY TWO FINGERS ARE 'BETTER' THAN TEN (and why ten fingers are 'better' than two)
and
WHY PLAIN ARE 'BETTER' THAN ROLLED (and why rolled fingers are 'better' than plain)

Comparing tables 1 and 3, one is immediately struck by the fact there are in the aggregate almost twice as many failures to identify with ten fingers as there are with only two. For each of the four data sets with ten fingers available (i.e., 1, 2, 3, and 4) in these studies, accuracy (TAR) with two fingers is higher than with ten.

Both the explanation for this phenomenon and the rationale for using ten fingers can be found in the filter rate. The BCII filter rate was 1.3% for ten fingers, 29.8% for two. Viewed another way, the ATB performed ($925 * 0.013 * 1.2M \cong$) 15M detailed matches to obtain a correct identification of 898 subjects (and no misidentifications) when using ten fingers, and performed ($925 * 0.298 * 1.2M \cong$) 330M detailed matches to obtain a correct identification of 913 subjects (and 11 misidentifications) when using two fingers. Using ten fingers, it incorrectly filtered (screened) out 27 mates in the process of avoiding 1095M matching operations, but made no misidentifications. Using two fingers, it incorrectly filtered out only 1 mate in the process of avoiding 780M matching operations, but made 12 misidentifications. Using two fingers, it used 22 times as many computer matching operations to increase the number of identifications by less than 2%.

A comparison of tables 9 and 11 reveals that the same phenomenon is exhibited by the other data sets.

Filtering, as performed on the ATB, uses some gross information (characterization) about all available fingers to reduce the number of detailed matches performed, and thus maintain the required throughput.

There are some direct consequences of this mode of operation of the ATB.

Plain images are smaller (ceteris paribus) than rolled images, and smaller images contain less information (cet. par.) than larger images. The less information presented to the characterization process, the fewer mates (incorrectly) filtered out before matching. Thus:
Provided that the fingerprints remain of sufficiently high quality (i.e., high 'matchability' by the matcher in question), the use of a reduced number of plain impressions rather than ten rolled impressions can be expected to result in a **modest increase in** the number of **correct identifications** (TAR) coupled with a **dramatic increase in** the **time** required to process each subject. Such a shift in the operating point of the system would require more resources to sustain the same throughput.

Studies of Plain-to-Rolled Fingerprint Matching

APPENDIX D:

ANALYSIS OF SOURCE MEDIUM OF IAFIS IMAGES

In an attempt to quantify the statement "inked images predominated", 100 records – properly, sets of images – of the 925 supplied from the FBI CMF were examined.

There are three caveats. First, the 925 records supplied by the FBI were not a random selection from the IAFIS CMF, but rather the CMF records for the subjects in the BCII set. Second, the 100 were not a random selection from the 925, but rather the first 100 in file name order (i.e., #100656MA0 through #180862RB8). Third, a set of 100 is rather small for the generation of precise statistics.

Visual examination showed that the 100 records came from the following sources:

> 50 cards, definitely inked: determination was made as follows:
>> 2 reversal of ridge & valley due to very high pressure squeezing out ink
>> 20 improper registration of images
>> 28 overlap of images (includes 1 also counted as a reversal)
>> 1 adhesive patch (covering up previous image)
> 14 cards, probably inked - cards with nothing conclusive to indicate that they were not inked [note that #144905AC5 might well have been a rescan, and that #142811CA4 seemed to be composed of images from different cards]
> 19 cards, printed from livescan, then rescanned
>> 12 Digital Biometrics, Inc (DBI) (of which 10 showed evidence of a separate placement of finger 10)
>> 7 Identix
> 17 livescan [note that images 2, 6, 11, & 12 of 130610WB8 were quite poor: artifacts from WSQ compression might be present]

Although one might argue that, knowing nothing else, the 14 should be allocated to ink and printed from livescan in proportion to the ratio of the 69 cards already so classified 'definitely inked' : 'printed from livescan' (that is, 10 inked, 4 printed from livescan), in fact all but one of the 14 presented the appearance of inked cards.

One may reasonably conclude that over 60% of the records came from inked cards, somewhat less than 20% came directly from livescan records, and about 20% came from rescanning cards that had been printed from livescan records.

Only one (#142811CA4) of the 100 records appeared to contain images from different sources; in the records from the Ohio CMF the presence of images from different sources was much more frequent.

Studies of Plain-to-Rolled Fingerprint Matching

APPENDIX E:

CONSOLIDATION

Fingerprint matchers are a subset of biometric matchers. Much of the analysis of biometric matchers employs generally accepted methodologies and terminologies, which are well described in the existing literature [4,10]. A necessary component of these is the concept of consolidation: bringing together into one record all images pertaining to a given subject. In the case of a system (such as the IAFIS or the ATB) which allows only one [active] record for each subject, that means purging all aditional records.

Consolidation is necessary to remove ambiguity and error from the process of scoring the matcher's accuracy.

The frequency of consolidations (of subjects with more than one record) in operational fingerprint databases is high enough to render almost meaningless measures of accuracy (e.g., FAR and TAR) unless the consolidations are identified and eliminated. The methodology for performing the elimination is not as important as its result: removing too many records is preferable to removing too few. The methods we employed to eliminate consolidations in the three sets of data forming group 2 are illustrative of some of the techiniques that can be used.

To address the problem of consolidation in the TXDPS data (set 2), we first seeded the background gallery with the rolled images of the gallery subset. We then created a virtual probe set from the gallery subset (by telling the ATB to treat the gallery subset as a probe set). Finally, we searched the entire seeded gallery with the virtual probe set. In effect, we probed both the original background portion of the gallery and the seeded portion of the gallery with the rolled images of the seeded gallery itself; any match was potentially an additional record of the subject from the seeded gallery set.

Ignoring matches of images to themselves, there were four possible classes of matches: probe set to seeded portion of gallery set (high score and medium score), and probe set to background portion of gallery set (high score and medium score).

We resolved both classes of probe set to seeded portion of gallery set matches by visually comparing the images; all of the high scores and most of the medium scores were in fact consolidations (the ATB does not report low scores when in normal operating mode).

The FBI resolved both classes of probe set to background gallery matches for the TXDPS data: we sent our images together with the FBI number of the reported match to the IAFIS facility in Clarksburg, WV. (We could not compare the images ourselves: we do not have the images of the 1.2M records in the background gallery, but rather only the feature vectors.) The FBI reported that all of the high scores were in fact consolidations, and also reported which of the medium scores were (or were not) consolidations.

For the remaining two sets, INS10 and INS_ben, we treated as a consolidation any high or medium score match to the background gallery. We were able to do so because of the proven reliability of the matcher itself as a detector of probable consolidations; we did so to avoid the additional time that would have been required by manual resolution.

We removed all records representing consolidations with the background gallery from further consideration. In data set 2, we allowed the record with the lowest ID number in a group of consolidated records in the gallery set to remain. Because we incrementally built and consolidated data sets 3 and 4, this process would have presented too much opportunity for error when adding more records to the set. Instead, in those two sets we removed from consideration **all** records associated with a consolidation.

Studies of Plain-to-Rolled Fingerprint Matching

APPENDIX F:

GLOSSARY

We assume in defining 'false accept', 'false reject', 'true accept', and 'true reject' that both query and target sets are fully consolidated.

AN2K   American National Standard ANSI/NIST-ITL 1-2000; a standard for the exchange of biometric data (cf EFTS) [13]

characterize, characterization – see feature vector

consolidation   for each subject in the [target | gallery | query | probe] set, the identification and elimination from that set of any additional fingerprint records for that subject; also, any such additional record; also, the very existence of any such additional record; see Appendix E

efts, EFTS   an efts record is a record compliant with the FBI's Electronic Fingerprint Transmission Specification [12]; EFTS is somewhat more restrictive than AN2K, so that an EFTS-compliant record is almost always compliant with AN2K; the ATB accepts efts records, but allows the omission of some mandatory fields [6]

enrollment   the entry of a subject's fingerprint record into the target set

false accept   the nomination as a match a subject that is not the mate; the false accept rate (FAR) is the rate w.r.t. number of probes of 'false accept' responses; false accept is defined only when the probe is an impostor

FAR – see false accept

false reject   the failure to nominate as a match the subject that is the mate; the false reject rate (FRR) is the rate w.r.t. number of probes of 'false reject' responses; false reject is defined only when the probe has a mate

FBI number   a six to nine character alphanumeric identifier assigned by the FBI to identify a subject; it is used (when known) as a definitive identifier by most law enforcement agencies in the United States; the existence of two records within IAFIS, with different FBI numbers but representing the same subject, would constitute a consolidation (qv)

feature vector   the ATB's extract of the matching information of the subject's images; the ATB denotes the process of extraction as 'characterization'

Studies of Plain-to-Rolled Fingerprint Matching

[to] filter   [verb] to create a subset of the original gallery with the goal of restricting matching to this subset, thus reducing the number of actual matches to be performed

finger position   a number used to denote a particular finger by its position on a tenprint card; 1 - right thumb, 2 - right index, ..., 6 - left thumb, ..., 10 - left little

fingerprint record   a set of fingerprint images of a single subject captured at a single session (e.g., the images on a tenprint card)

FRR – see false reject

gallery [set]   a set of fingerprint records of known subjects against which a search is conducted; a consolidated gallery is a gallery with exactly one record per subject; a gallery set is a (possibly improper) subset of the target set [10]

identification   the determination of which subject (if any) in the gallery corresponds to the current subject of interest in the probe; also referred to as 1-to-n or one-to-many matching

impostor   a probe that has no mate in the gallery

impression   an image, whether using ink or a 'livescan' optical device, of a finger; a plain (or 'flat') impression is made by pressing the bottom surface of the terminal phalange of the finger or thumb against a flat surface, whereas a rolled impression is made by rotating the finger or thumb around its longitudinal axis without slip while causing the same two surfaces to bear

match   an indication from the matching system that a particular subject within the gallery bears a notable similarity to the subject in question from the probe; also, the particular subject in the gallery

match score   a numeric score showing the degree of similarity in a match; generally presumed to be a non-negative increasing monotonic function

mate   the subject in the gallery corresponding to the subject in the probe (or vice versa)

probe [set]   a set of fingerprint records of unknown subjects to be identified or verified that is used in performing a search; a consolidated probe is a probe with exactly one record per subject; a probe set is a (possibly improper) subset of the query set [10]

query [set]	the set of fingerprint records of unknown subjects to be identified or verified; a consolidated query is a query with exactly one record per subject [10]
rank-1 match	the match with the score showing the greatest degree of similarity when more than one match is reported for a member of the probe set; otherwise, the match, if any [in analyzing the ATB results, the rank-1 match was treated as the only reported match]
registration	the placement of a tenprint card on a scanner such that the fingerprint image areas on the card physically correspond to the scanner's construct of them

repository – see target set

[to] seed	[verb] given a gallery set consolidated with respect to the probe set (that is, with (presumably) no mate to any subject in the probe set), to enroll into the gallery exactly one mate for each subject in the probe
segmentation	the resolution of a multi-finger (e.g., slap-4) image into its constituent fingerprint images
slap-4	a set of four fingerprints – the four fingers of either hand – taken simultaneously; necessarily these are plain impressions rather than rolled
submission	the presentation to the ATB of a probe; the images so submitted
target [set]	the set of fingerprint records of known subjects; a consolidated target is a target with exactly one record per subject; to the ATB, the target is also known as the repository [6][10]
tenprint card	a card, such as FBI forms FD-249 or FD-258, intended to capture a fingerprint record; the exemplars have locations for fourteen images: ten for a rolled impression of each digit, two for a plain impression of each thumb, and two for a plain impression of the four fingers (taken simultaneously) of each hand; the latter two are referred to as slap-4 images
true accept	the nomination as a match a subject that is the mate; the true accept rate (TAR) is the rate w.r.t. number of probes of 'true accept' responses; true accept is defined only when the probe has a mate

TAR – see true accept

true reject the nomination of no subject as a match when in fact there truly is no mate; the true reject rate (TRR) is the rate w.r.t. number of probes of 'true reject' responses; true reject is defined only when the probe is an impostor

TRR – see true reject

verification the process of testing whether the query image is a mate of the target image; also referred to as one-to-one matching; note that the ATB's design emphasizes its primary role as an identification (vs. verification) matcher

APPENDIX G:

NOTATION in black and white

P	probe set (any)	$\{p_i\}$
P^I	probe set (impostor)	$\{p_i: \forall p_i \in P^I \ (!\exists g_k \in G \ni p_i \ \mathbf{M} \ g_k)\}$

in English: there is no mate

P^M	probe set (mated)	$\{p_i: \forall p_i \in P^M \ (\exists g_k \in G \ni p_i \ \mathbf{M} \ g_k)\}$

in English: there is a mate

$\mathbf{M}$ ate: $\qquad$ ($p_i \ \mathbf{M} \ g_k$) iff $\exists$ a subject $s_j \ni$ both p_i and g_k are images of fingerprints taken from s_j [thus $\mathbf{M}$ is transitive and commutative]

$\mathbf{R}$ eported Match: ($p_i \ \mathbf{R} \ g_k$) iff matcher reports some g_k as the rank-1 match for p_i

G	gallery set (any)	$\{g_i\}$
G^C	gallery set (consolidated)	$\{g_i: \forall g_i \in G^C \ (!\exists g_{k \ i} \in G^C \ni g_i \ \mathbf{M} \ g_k)\}$
	or	$\{g_i: \forall g_i \in G^C \ (\forall g_{k \ i} \in G^C \ !(g_i \ \mathbf{M} \ g_k))\}$

in English: there is no mate within the gallery

From the definitions of P^M, G^C, and $\mathbf{M}$:

$\qquad \forall p_i \in P^M \ (\subset \{g_k \in G^C \ni p_i \ \mathbf{M} \ g_k\} = 1)$

in English: each member of the mated probe set has exactly 1 mate in the consolidated gallery

$\qquad P^I \cup P^M = P$

$\qquad P^I \cap P^M = M$

FAR, FRR, TAR, TRR:

False Accept:	($p_i \ \mathbf{R} \ g_k$) $\wedge$!($p_i \ \mathbf{M} \ g_k$) $\wedge$ $p_i \in P^I$
False Reject:	!($p_i \ \mathbf{R} \ g_k$) $\wedge$ ($p_i \ \mathbf{M} \ g_k$) $\wedge$ $p_i \in P^M$
True Accept:	($p_i \ \mathbf{R} \ g_k$) $\wedge$ ($p_i \ \mathbf{M} \ g_k$) $\wedge$ $p_i \in P^M$
True Reject:	!($p_i \ \mathbf{R} \ g_k$) $\wedge$!($p_i \ \mathbf{M} \ g_k$) $\wedge$ $p_i \in P^I$

FAR = $\subset$ {False Accept} / $\subset$ (P^I)

FRR = c {False Reject} / c (P^M)
TAR = c {True Accept} / c (P^M)
TRR = c {True Reject} / c (P^I)

Note that for each probe there is exactly one result: False Accept $|$ False Reject $|$ True Accept $|$ True Reject. However, each probe is in exactly one of P^I or P^M.
For those in P^I, only FAR and TRR are possible (defined). Similarly, fo$|$r those in P^M, only TAR and FRR are possible. Thus FAR + TRR = 1, and TAR + FRR = 1. Hence, FAR + TRR + TAR + FRR = 2.

April 2004