AD-A056 918 MITRE CORP BEDFORD MASS F/G 15/3

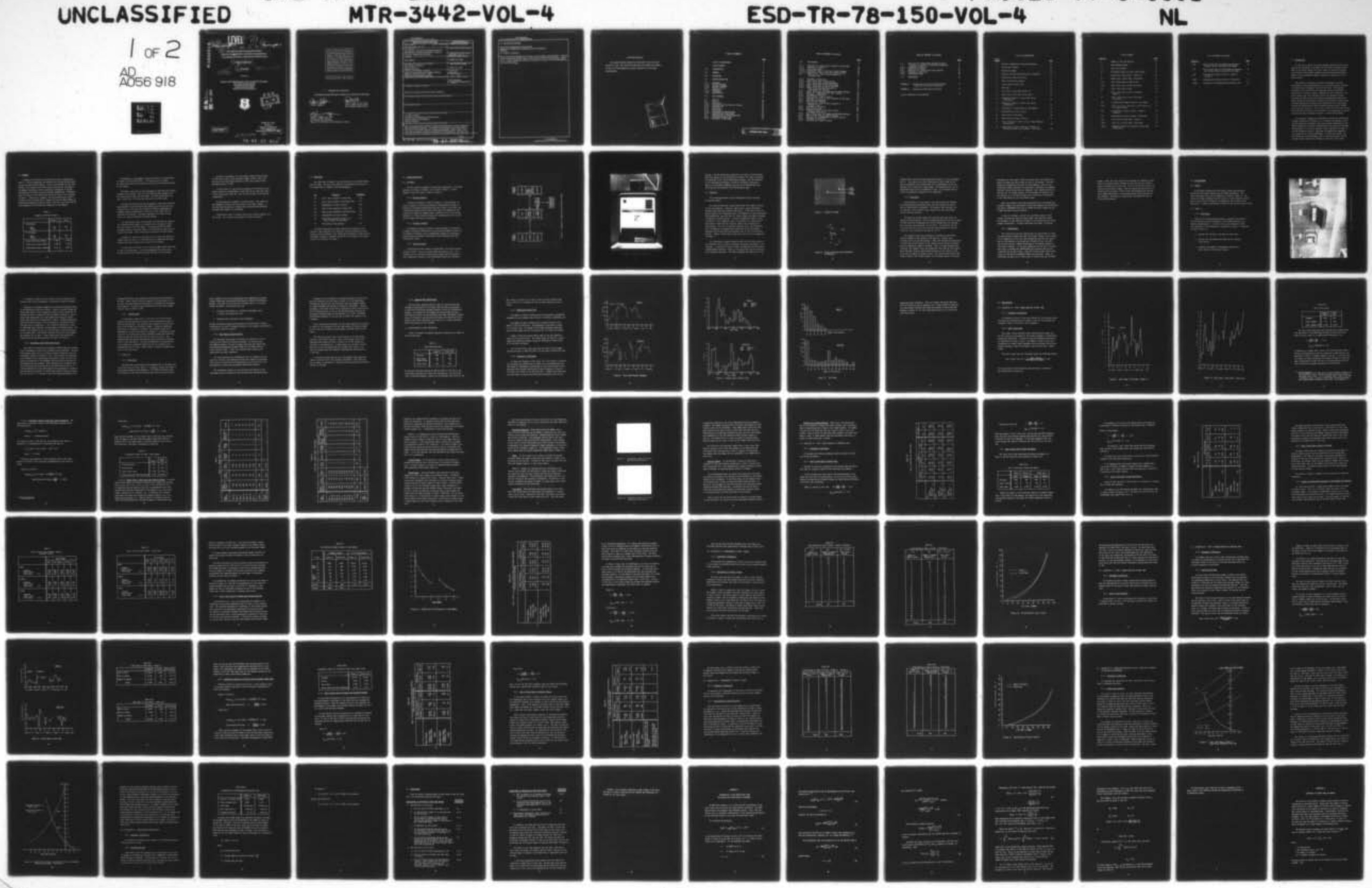
TEST RESULTS-ADVANCED DEVELOPMENT MODELS OF BISS IDENTITY VERIF--ETC(U)

JUL 78 P BENSON F19628-77-C-0001

UNCLASSIFIED MTR-3442-VOL-4 ESD-TR-78-150-VOL-4 NL

1 OF 2

AD A056 918



ESD-TR-78-150, Vol [redacted] -4

MTR-3442, Vol [redacted] -4

TEST RESULTS-ADVANCED DEVELOPMENT MODELS OF BISS IDENTITY VERIFICATION EQUIPMENT, VOLUME IV, AUTOMATIC FINGERPRINT VERIFICATION,

BY PETER BENSON

JULY 1978

Technical rept.,

98p.

Prepared for

DEPUTY FOR SURVEILLANCE AND NAVIGATION SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
Hanscom Air Force Base, Massachusetts

Project No. 4130

Prepared by

THE MITRE CORPORATION
Bedford, Massachusetts

Contract No. F19628-77-C-0001

REVIEW AND APPROVAL

This technical report has been reviewed and is approved for publication.

THOMAS F. BASHARA
Project Manager

PAUL E. PEKO, Lt Col, USAF
Chief, Engineering Division
BIS Systems Program Office

ROGER T. KOZUMA, Col, USAF
Systems Program Director
BIS Systems Program Office
Deputy for Surveillance and Navigation Systems

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-78-150, Vol. IV	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Test Results — Advanced Development Models of BISS Identity Verification Equipment, Volume IV, Automatic Fingerprint Verification		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER MTR-3442, Vol. IV
7. AUTHOR(s) Peter Benson		8. CONTRACT OR GRANT NUMBER(s) F19628-77-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS The MITRE Corporation P.O. Box 208 Bedford, MA 01730		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project No. 4130
11. CONTROLLING OFFICE NAME AND ADDRESS Deputy for Surveillance and Navigation Systems Electronic Systems Division, AFSC Hanscom Air Force Base, MA 01731		12. REPORT DATE JULY 1978 13. NUMBER OF PAGES 96
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ACCESS CONTROL
AUTOMATIC FINGERPRINT VERIFICATION
DIGITAL SIGNAL PROCESSING
ENTRY CONTROL

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This volume presents the results of testing the developmental model for fingerprint personal identity verification. The purpose of the program was to determine the Type I error rate (false rejection of authorized personnel), random Type II error rate (false admittance of unauthorized personnel), and throughput. The tests were conducted both at the Verification Laboratory at The MITRE Corporation and in the — over

(over)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

78 07 27 028

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19. Key Words (continued)

MINUTIAE/FINGERPRINT RECOGNITION
PERSONAL ATTRIBUTE AUTHENTICATION TECHNIQUES
TESTING

20. Abstract (continued)

field at the Weapons Storage Area at Pease Air Force Base, New Hampshire. Volumes II and III present the detailed results of the speaker and handwriting systems, respectively. Volume I presents an executive summary and Volume V several miscellaneous but related subjects.

ACKNOWLEDGMENT

This report has been prepared by The MITRE Corporation under Project No. 4130. The contract is sponsored by the Electronic Systems Division, Air Force Systems Command, Hanscom Air Force Base, Massachusetts.

TABLE OF CONTENTS

PRECEDING PAGE BLANK

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (concluded)

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF TABLES (concluded)

1.0 INTRODUCTION

As part of its effort to develop external physical security systems for the Department of Defense, the Electronic Systems Division (ESD) of the United States Air Force, under its Base and Installation Security System (BISS) program, acquired advanced development models of three automated identity verification systems for use in entry control.

The identity verification systems are designed to provide verification of the claimed identity of persons entering or leaving a restricted area through an Entry Control Point. The systems accomplish this verification by examining unique and measurable properties of a person's speech, handwriting, and fingerprint. The three systems that were evaluated include the Automatic Speaker Verification (ASV) system, the Automatic Handwriting Verification (AHV) system and the Automatic Fingerprint Verification (AFV) system. The MITRE Corporation was given the responsibility for installing, testing and evaluating the three systems/techniques both at MITRE and at a field test site. This volume, Volume IV, of the test report will address in detail the tests, results, and evaluations of the AFV system as briefly discussed in Volume I, the Executive Summary.

The Automatic Fingerprint Verification system was acquired from Calspan Corporation of Buffalo, N.Y., by ESD in the Spring of 1976. The AFV system was installed in the Entry Control Laboratory at The MITRE Corporation in Bedford, Massachusetts and tested and evaluated during Phase II of the test program. After completion of Phase II, the AFV system was taken to Pease AFB, Portsmouth, New Hampshire, for evaluation in a field environment. In both test phases, the AFV system was tested alongside, and in conjunction with, the AHV and ASV systems described in Volumes II and III of this report.

2.0 SUMMARY

A summary of the Phase II and Field Test data is presented in Table I. The AFV system was not available for testing during Phase I as were the ASV and AHV systems. Phase II was divided into two test periods. The results of the first test period are not included in the summary because, as will be reported, a parameter of the AFV algorithm was in slight error and required modification. Therefore, only the results of the second test period of Phase II are included in Table I. There was no change in the algorithm between Period 2 of Phase II and the Field Test. Because of the very small number of females participating in the Field Test, only the error rates for the male population are included for comparison in the summary.

Table I

Summary of AFV Test Results

	Phase II	Field
Enrollments		
Male	166	268
Female	34	7
Participants		
Male	162	202
Female	34	6
Type I Error Rate (Male)	4.58%	6.54%
Type II Error Rate (Male)	2.18%	2.32%
Verification Time (seconds)	8.53	8.94

Enrollments are the number of people enrolled in the AFV system for Phase II and the Field Test. Participants are the number of people that used the AFV system at least once after being enrolled in the system.

The Type I error rate is the percentage of users who were falsely rejected by the system. That is, they were who they claimed to be but they could not pass the AFV system. The BISS requirement of a Type I error rate of 1% was not achieved by the AFV system.

The Type II error rate is the percentage of users who could have been accepted by the AFV system even though they were not who they claimed to be. This data was obtained by computer comparisons of randomly selected fingerprints, collected in real time, with the reference file library. The BISS requirement of a Type II error rate of 2% was not achieved by the AFV system.

The verification time is the average time from when an individual's identification number is accepted until a decision is made for that person by the AFV system. To provide the BISS required throughput using a detention module, the verification time could not exceed six seconds. This time was not achieved by the AFV system.

In Phase II, the Type I and Type II error rates were higher for females than for males. In the Field Test, no comparison was made because only seven females were enrolled on the system.

In both tests, the Type I error rate was higher before noon than it was in the afternoon. It has been suggested that this occurs because the hand is relatively dry and oil-free in the morning.

The Type I performance for the Primary finger was much better (≈14%) than it was for the Alternate finger (≈43%), and it was much better for the Primary matcher (≈15%) than it was for the Secondary matcher (≈53%).

A large percent of the enrolled population had difficulty verifying in spite of reenrollments. About 22% of the users had a Type I error rate greater than 5%. Most of this was due to poor quality fingerprints.

The AFV system was awkward for several users. The scanner is not designed for those with very large or very small fingers. A few users with arthritis or unusual bone structure had difficulty placing their fingers in the scanner.

Sophisticated Type II spoofing tests were conducted against the AFV system. The results are presented in another document.

3.0 OBJECTIVES

The objectives of Phase II and the Field Test for the AFV system were the same. The objectives that will be discussed in Section 6.0 are listed below by number, title and paragraph.

No.	Objective	Paragraph
23	Type I Error Analysis in Real Time	6.1
24	Type I Error Analysis in Non-Real Time	6.2
25	Independence of Type I Scores	6.3
26	Type II Error Analysis in Real Time	6.4
27	Type II Error Analysis in Non-Real Time	6.5
28	Independence of Type II Scores	6.6
29	Sensitivity Analysis of Type I and Type II Errors to Thresholds	6.7
30	Verification Time Analysis	6.8

A detailed discussion of these objectives including the test approach, data required, data reduction, and data analysis are presented in other documents. The objectives pertaining to human factors, enrollment procedure, equipment reliability, environment, and hybrid systems are discussed in Volume V.

4.0 SYSTEM DESCRIPTION

4.1 EQUIPMENT

The AFV system is composed of three basic components: an Entrant Terminal, an Operator Terminal, and a Central Station. A block diagram of the AFV system is shown in Figure 1.

4.1.1 Entrant Terminal

The Entrant Terminal is shown in Figure 2. This terminal contains a keypad for entering the entrant's identification number, a display for providing instructions and entry status to the user, and a mechanical-electro-optical sensor that scans the fingerprint and converts the image into digital data. Mounted on top of the entrant terminal is a finger position display for assisting the user in placing his finger correctly in the scanner.

4.1.2 Operator Terminal

The Operator Terminal contains a similar keypad to the Entrant Terminal, functional pushbuttons for enrolling and general control of the system's operation, and an alphanumeric display for displaying system status, enrollment instructions, alarm conditions and user verification status.

4.1.3 Central Station

The Central Station contains a preprocessor, the central processing unit (CPU) (a Texas Instrument 980B minicomputer) and a disc storage unit for storing the application program and the library of user fingerprint reference files established during the enrollment

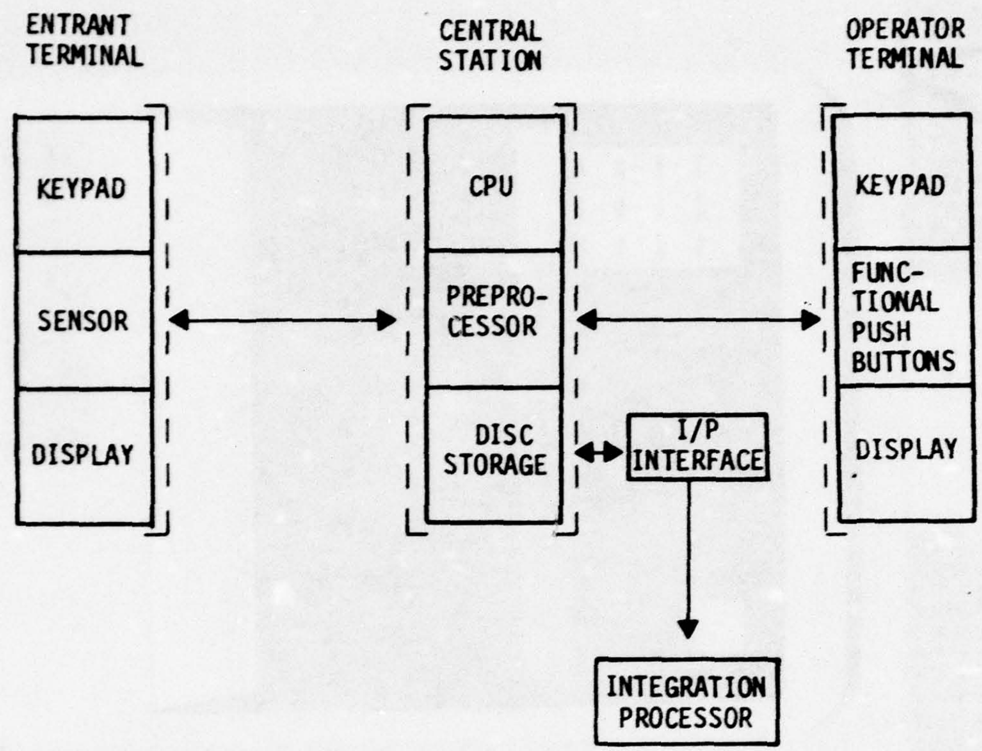


Figure 1. Automatic Fingerprint Verification System.

Figure 2. AFV Entrant Terminal.

process. The AFV system was connected to the CPU of the ASV system which was employed as an Integration Processor (IP) that coordinated the operation of all three verification systems. The CPU forwarded to the Integration Processor, via the IP interface, appropriate data to be recorded on magnetic tape for use in determining error rates and for the printout of AFV verification data.

4.2 OPERATION

The following provides a brief description of the operation of the AFV system.

The scanner mechanism, located under where the finger is placed on the user terminal, contains a light source, prism, photodiode detector, and associated mirrors and drives. A 3/8 inch by 3/8 inch binary image of the fingerprint area is transmitted to the preprocessor. The preprocessor scans the data for ridge endings and ridge branchings collectively called minutiae. A typical fingerprint image is shown in Figure 3. The fingerprint data is comprised of a collection of unit vectors which show the relative position and orientation of the minutiae. This set of minutiae position ($X_i Y_i$) and orientation (θ_i) coordinates is compared with coordinates of an individual's minutiae data obtained during enrollment and on file in the Central Station. Figure 4 shows many of the unit vectors for the fingerprint of Figure 3.

In correlating a person's minutiae data with the data stored in the reference file, a score is produced indicating the degree of correlation. A pass/fail decision is made by comparing the score with a predetermined threshold. There are two algorithms used for determining a pass/fail decision. The first algorithm is referred to as the

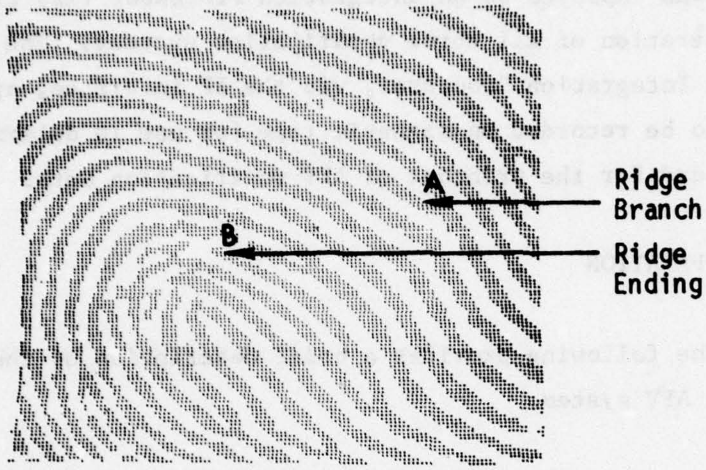


Figure 3. Fingerprint Image

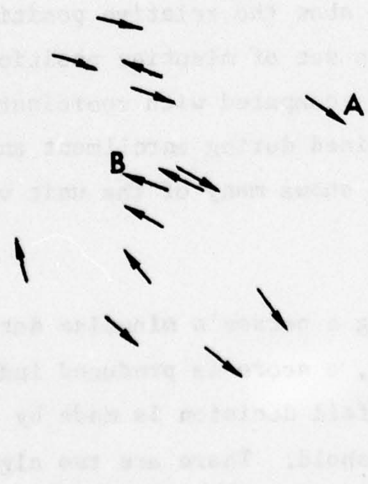


Figure 4. Relative Position and Orientation of Minutiae.

Primary matcher and the second algorithm referred to as the Secondary matcher. The choice as to which algorithm is used is made at the time of enrollment and is determined by the performance of the two matchers with a particular finger image. The Primary matcher is the preferred matcher and is predominantly used. The algorithm that is chosen is permanently assigned to the individual.

4.2.1 Enrollment

Upon arriving at a test area, a test participant was informed by an audiovisual presentation about the test program and enrollment and verification procedures. The enrollee then provided, via an enrollment card, personnel data required for the verification process and data analysis.

To enroll in the AFV system, the enrollee must have first enrolled in the ASV system. This was required because of the way in which the AFV system was integrated into the IP configuration for testing, data recording and generation of statistics. The personal data entered into the CPU of the ASV system was made available to the AFV system for the recording of AFV data by the IP.

The operator began AFV enrollment by entering the enrollee's identity number via the keypad on the Operator Terminal thus creating a file in the AFV computer. Next, the enrollee was instructed to place each of his fingers in the scanner (on the Entrant Terminal) where a fingerprint image is obtained. After all eight fingers were observed, the computer chose the fingerprint that best met the requirements through a Quality Index based primarily on the clarity of the print and the number of minutiae found. The enrollee then placed that finger, known as the Primary finger, in the scanner where it was held stationary for four scans. After the fourth scan, the enrollee removed his finger and then placed it back in the scanner. This procedure was repeated for the next six scans making a total of ten fingerprint images available for generating a reference file against which all future prints would be compared. Upon completing the enrollment for the Primary finger, a second finger was chosen by the AFV computer as an Alternate finger and a reference file generated in the same manner as for the Primary finger.

After the reference files were established for both Primary and Alternate fingers, the enrollment process was complete and the enrollee was given a FINGERSCAN user card to remind him which finger was his Primary and which was his Alternate.

It was not uncommon to repeat the enrollment process either because the enrollee used the wrong finger, had difficulty using the terminal, or the computer, because of poor matcher scores, would suggest reenrollment. A normal enrollment took about 3.0 minutes.

4.2.2 Verification

The verification process required the test participant to enter his identification number and place his Primary finger in the scanner. If the comparison of the real time fingerprint data with the user's data in the reference file satisfied a threshold level, the user was verified and got a THANK YOU message on the message display of the entrant terminal. If the data comparison did not pass the threshold level, a message TRY FINGER AGAIN was displayed and the user reinserted his Primary finger. When the Primary finger failed three scans, the message TRY ALTERNATE FINGER was displayed. Again the individual was given up to three scans to verify using his Alternate finger. Thus, two basic three scan strategies are combined to generate a six scan decision strategy and the user was not rejected unless he failed on all six scans: three on the Primary finger and three on the Alternate finger. When an authorized person failed to verify on all six scans, a Type I error was generated and the user got a THANK YOU message on the message display indicating that he was finished. Detaining or rejecting the individual was not part of the test procedure.

5.0 TEST PROGRAMS

5.1 GENERAL

The Automatic Fingerprint Verification system was evaluated during two test phases conducted by The MITRE Corporation. The first test phase took place at MITRE and is known as Phase II. The second test phase evaluated the AFV system in a more realistic environment at Pease AFB in Portsmouth, New Hampshire and is known as the Field Test. A general description of Phase II and the Field Test follows.

5.2 PHASE II

5.2.1 Description

Phase II was conducted during August, September and October of 1976 at the Entry Control Laboratory of The MITRE Corporation in Bedford, Massachusetts. A picture of the user terminals as they were installed in the laboratory is presented in Figure 5. The goals of Phase II were to:

1. Evaluate the AFV Type I and Type II error rates.

2. Observe user performance and identify any learning curve effects.

3. Establish the degree of independence between the three identity verification systems.

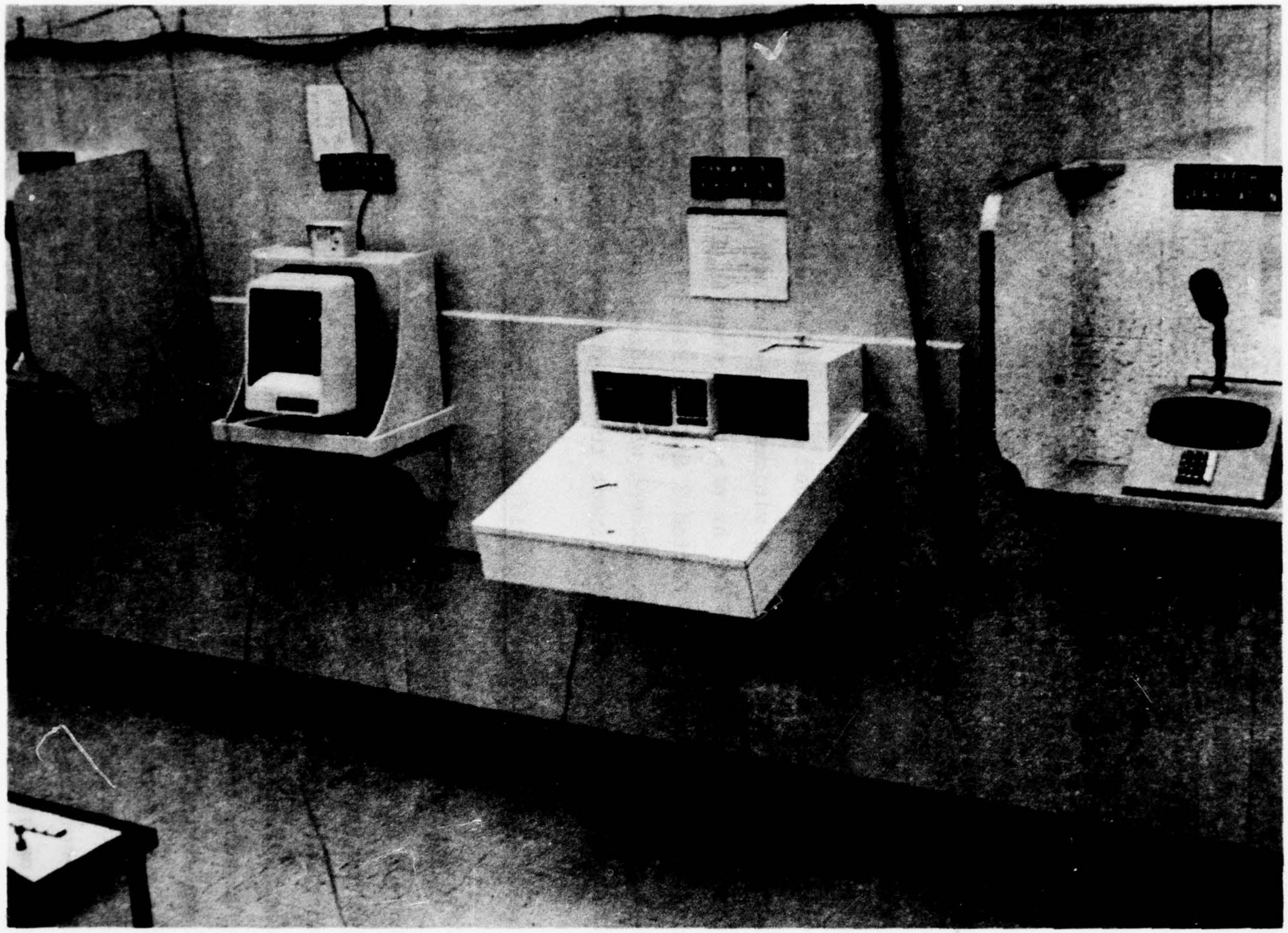


Figure 5. MITRE Entry Control Laboratory

In addition, Phase II was to satisfy the test objectives and provide other related information as detailed in another document.

Because of a parameter change in the decision algorithm, Phase II was divided into two test periods. Period 1 occurred during the first three weeks of test followed by Period 2 which lasted the remaining six weeks. During the first few weeks of Phase II, Type II errors were unexpectedly low, and individuals with good quality prints were found to be causing Type I errors. Calspan Corporation found the reason for this to be a parameter of the decision logic that was set to a wrong value. For a pass/fail decision to occur, an individual must register a minimum number of minutiae pairs. For Period 1 this number was set at six pairs when it should have been set at three pairs. For Period 2 and the Field Test, the parameter was properly set to three. Except where noted, this report will discuss only the results pertinent to Period 2 of Phase II.

5.2.2 Enrollment and Verification Process

The Phase II population was composed of MITRE employees consisting of programmers, technicians, engineers and secretaries. The test participants would arrive at the Entry Control Laboratory as per an enrollment schedule in groups of four. Upon arrival, the enrollees were informed, via an audiovisual presentation, about the test program, the operation of the equipment and what would be expected of them during the test program. After completing an enrollment card, the test participant was enrolled in the AFV system as per 4.2.1. After enrollment, the test participants were encouraged to return to the laboratory twice daily to verify their identity using the equipment as per 4.2.2. To prevent a user from being discouraged if he consistently failed, the verification results were not presented to the user. When a person failed to appear at the laboratory on any one day, the system operator contacted the individual and encouraged participation. The Entry Control Laboratory was open each weekday from the hours of 0800 to 1630.

5.2.3 Imposter Data

To establish a random Type II error rate for the AFV system, imposter data was gathered on every fourth verification attempt of each test participant. The imposter data consisted of three scans each of the Primary and Alternate fingers. Unless an individual kept track of the number of attempts they had made, they had no knowledge as to when they would be an imposter. An imposter attempt could not be differentiated from an attempt in which the individual failed to pass the verification on the first five scans. The imposter data was used by the computer during off hours to generate Type II error statistics. This was accomplished by comparing each of the imposter's fingerprint data with all the reference file fingerprint data of others. When a match occurred, the appropriate Type II error was noted.

5.3 FIELD TEST

5.3.1 Description

The Field Test was conducted from November 1976 to February 1977 at the Entry Control Point (ECP) to the Weapons Storage Area (WSA) at Pease AFB, Portsmouth, New Hampshire. The MITRE Corporation under the direction of the Base and Installation Security System Program

Office (BISSPO) acted as the Responsible Test Organization in supervising, installing, testing and evaluating the Identity Verification Systems. The goals of the Field Test included those in 5.2.1 plus several new goals. The new goals were to:

1. Determine performance in a realistic environment with a typical user population, and

2. Determine user reactions to the techniques.

Further information regarding the details of the Field Site, overall system set-up and other information related to the purpose and objectives is provided in another document.

5.3.2 Enrollment and Verification

All personnel with access to the WSA were scheduled to be enrolled in the Identity Verification Systems. Of the more than 400 badged personnel, the most frequent users were the people from the Munitions Maintenance Squadron (MMS) and the 509th Security Police Squadron (SPS). Other less frequent users worked in the Engineering and Supply Squadrons.

The user population was predominantly male (6 females) and consisted of officers, non-commissioned officers, airmen and civilians. Their jobs were varied and included mechanics, security police, administrators, electricians and general maintenance workers.

The enrollment process for the Field Test was similar to the enrollment process of Phase II with the exceptions described herein.

Because of the environment of the user population for the Field Test was different from the users in Phase II, a new audiovisual briefing was designed and the enrollment card was changed slightly to reflect personal data unique to a military environment. During enrollment, there were two Air Force operators on duty in the van and one Air Force Operator on duty at the briefing which was given at the guard building located at the entrance to the WSA. The MITRE Test Engineer remained in the test van to ensure the quality of enrollments.

After completing enrollment, the test participants were required to verify on the equipment each time they entered or exited the WSA. The test van was in operation each weekday from 0600 hours to 1630 hours.

The test participants were never detained if the AFV system failed to verify their identity. When a participant failed, his identification number and time of failure on the AFV system was printed on a teletypewriter. In many cases, the operators on duty were able to ask the participant why he might have failed. The failure, reasons given and operator's observations were entered in a log book.

It should be noted that the test participants were required to adhere to the Air Force's own security procedures. The Field Test was not meant to supplement or interfere with the security procedures normal for the WSA.

5.3.3 Imposter Data (Field Test)

The Field Test imposter data for Type II error analysis was obtained as described in 5.2.3 except that the data was gathered on every 8th verification attempt instead of every 4th verification attempt. The reason for this change was because there was more verification activity in the Field Test with some individuals verifying up to 8 times in one day and also because the collection of the imposter data takes additional time, which, for a large population, would cause queuing at the terminal.

5.4 PARTICIPATION OF TEST POPULATIONS

Table II presents the general population statistics for Phase II and the Field Test.

Table II

User Population Data

	PHASE II	FIELD TEST
Enrolled	200	275
Male Users	162	202
Female Users	34	6
Total Users	196	208

Of the total enrolled population, 98% for Phase II and 76% for the Field Test returned for at least one verification. For the Field Test it had been planned to enroll all 450 persons with access to the

WSA; however, because of the time of year and work schedules many people would not be visiting the area so those people were not enrolled.

5.4.1 Pedestrian Traffic Flow

The number of people verifying each day fluctuated considerably depending on work schedule and proximity to holidays, snowstorms, etc.

The number of daily verifications is shown in Figure 6 for Phase II and the Field Test. The beginning of each week is shown by a vertical line. The breaks in the graph represent those days when the AFV system was not in operation due to storms or holidays.

Figure 7 shows the typical daily usage of the AFV system for Phase II and the Field Test. The figure shows for Phase II a heavy usage that gradually declines during the mornings but again picks up in the early afternoon.

For the Field Test, there were three periods of heavy usage: between the hours of 0640 and 0740, 1120 and 1200, and 1540 and 1640.

5.4.2 Frequency of AFV Usage

To show the frequency of AFV usage, the total number of verification attempts was divided into 14 groups of ten verifications each. The number of individuals contributing to the verifications in each group was then collated and plotted in the histograms of Figure 8 for Period 2 of Phase II and the Field Test. Though there were about the same number of users, the Field Test particpants used the AFV

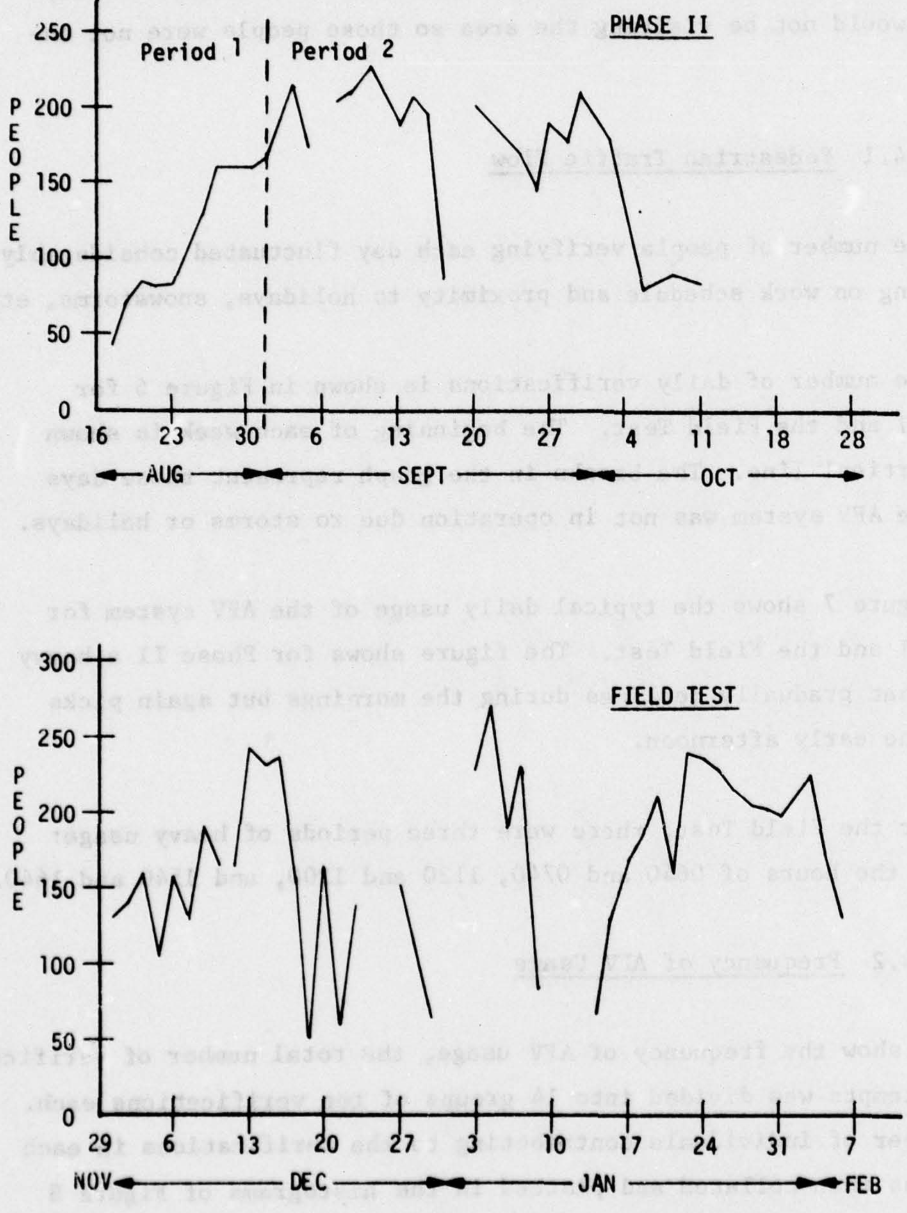


Figure 6. Daily Verification Attempts.

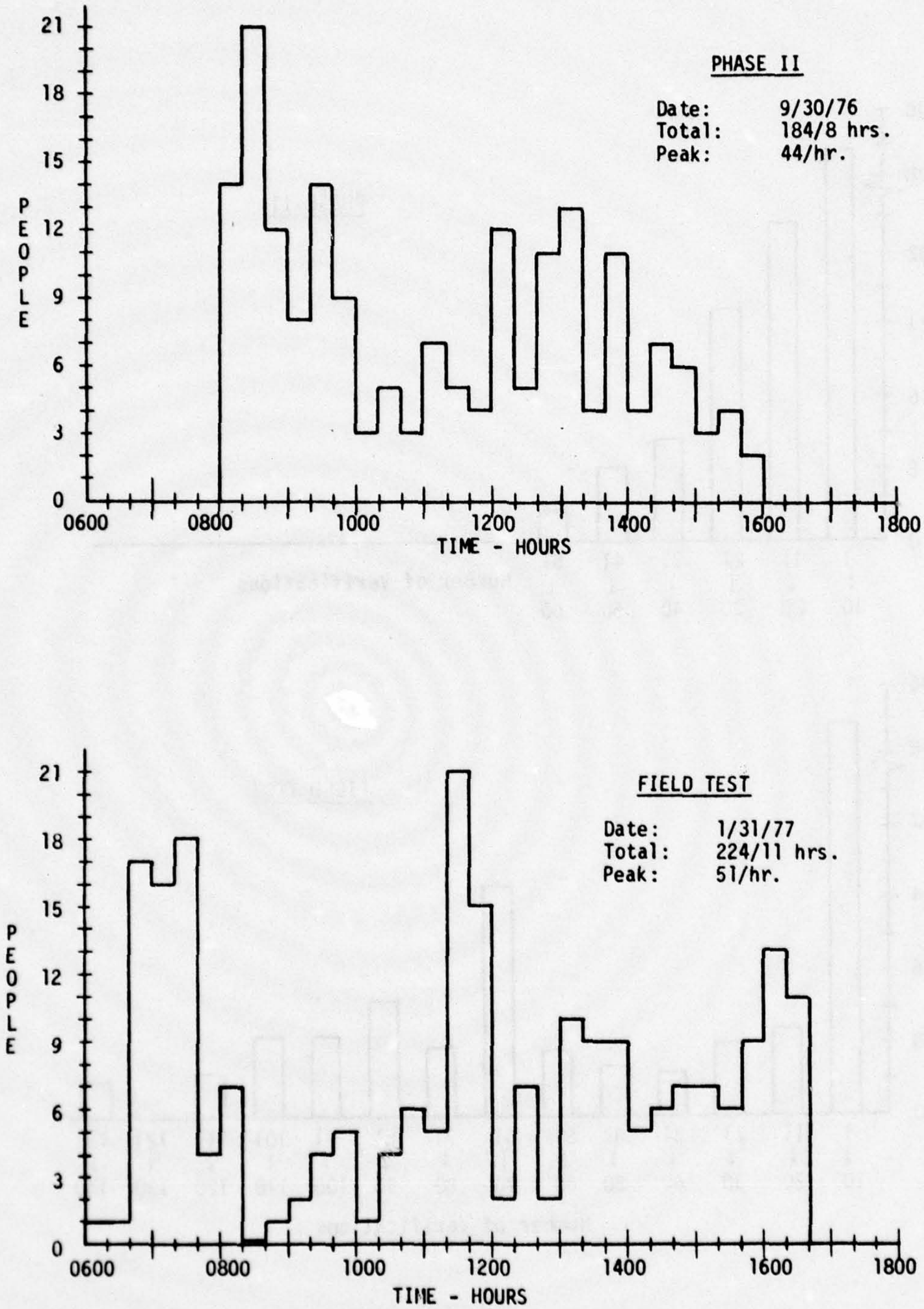


Figure 7. Typical Daily Traffic Flow.

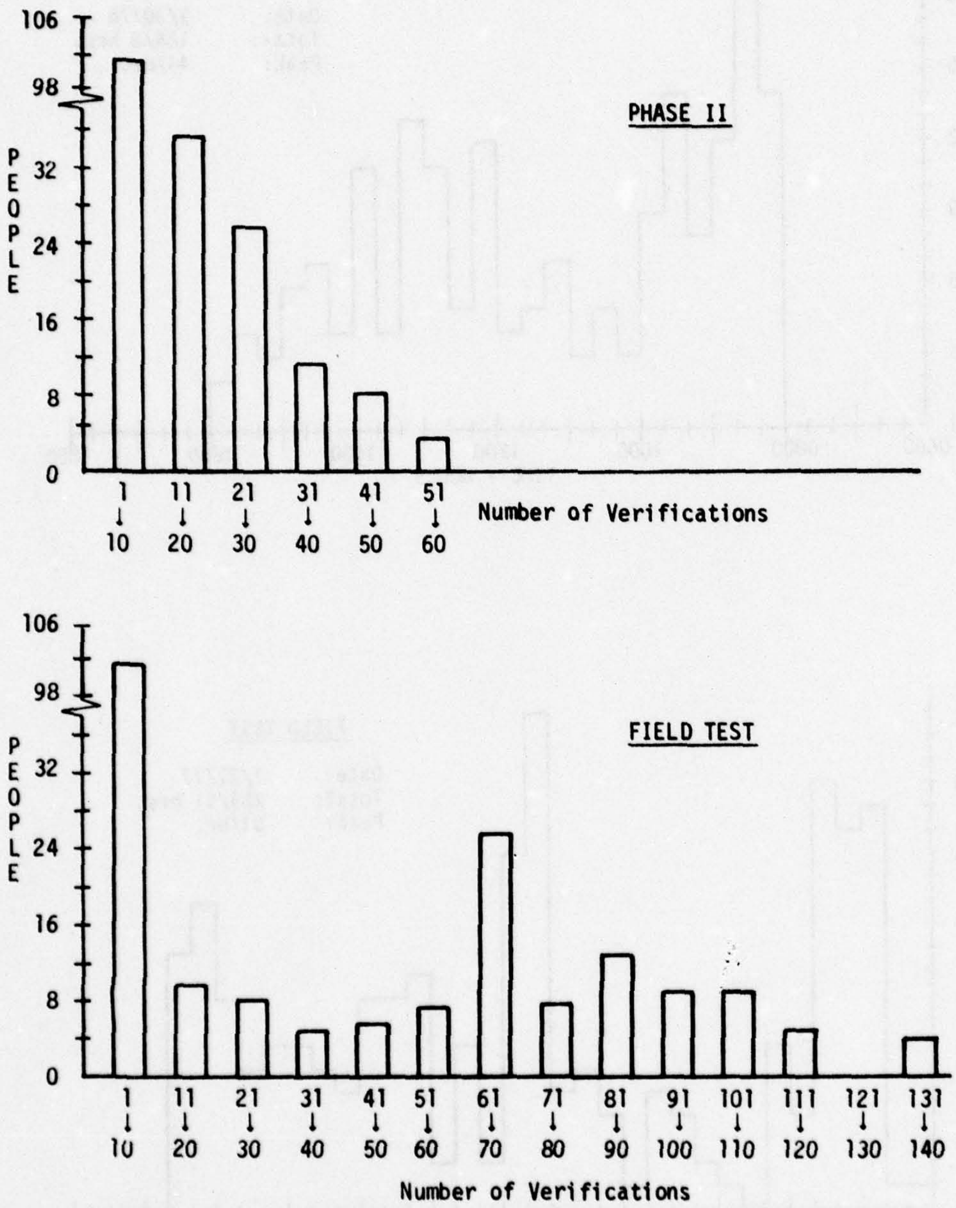


Figure 8. AFV Usage.

system much more frequently. This, of course, was due to the fact that the Field Test was in operation for a longer period of time and the participants verified each time they entered or exited the WSA. Whereas, in Phase II, the test participants were not to verify more than twice daily.

6.0 TEST RESULTS

6.1 OBJECTIVE 23 - TYPE I ERROR ANALYSIS IN REAL TIME

6.1.1 Statement of Objective

To estimate the Type I error rate (failure to verify proper identity when the representation is actually true) and to determine the confidence limits of the Type I error estimate.

6.1.2 Type I Error Rate

Each time a person entered a valid identification number into the AFV system via the keypad on the Entrant Terminal, a verification attempt was registered. A failure to pass an attempted verification is classified as a Type I error. The number of Type I errors were accumulated on a daily basis throughout the test periods and are presented in Figures 9 and 10 for Phase II and the Field Test respectively.

The Type I error rate was determined using the following formula:

$$\text{Type I Error Rate (\%)} = \frac{\text{Type I Errors}}{\text{Verification Attempts}} \times 100$$

The error rates for the Field Test and for Period 2 of Phase II are listed below in Table III.

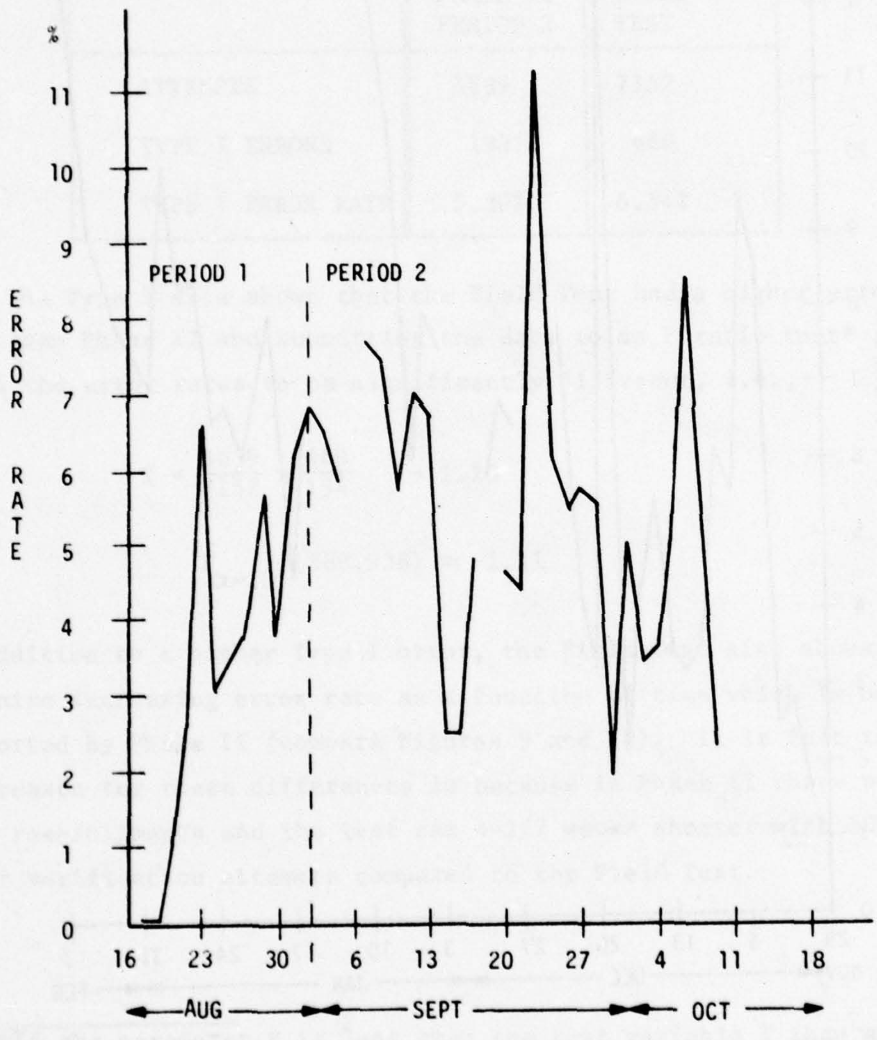


Figure 9. Daily Type I Error Rate - Phase II.

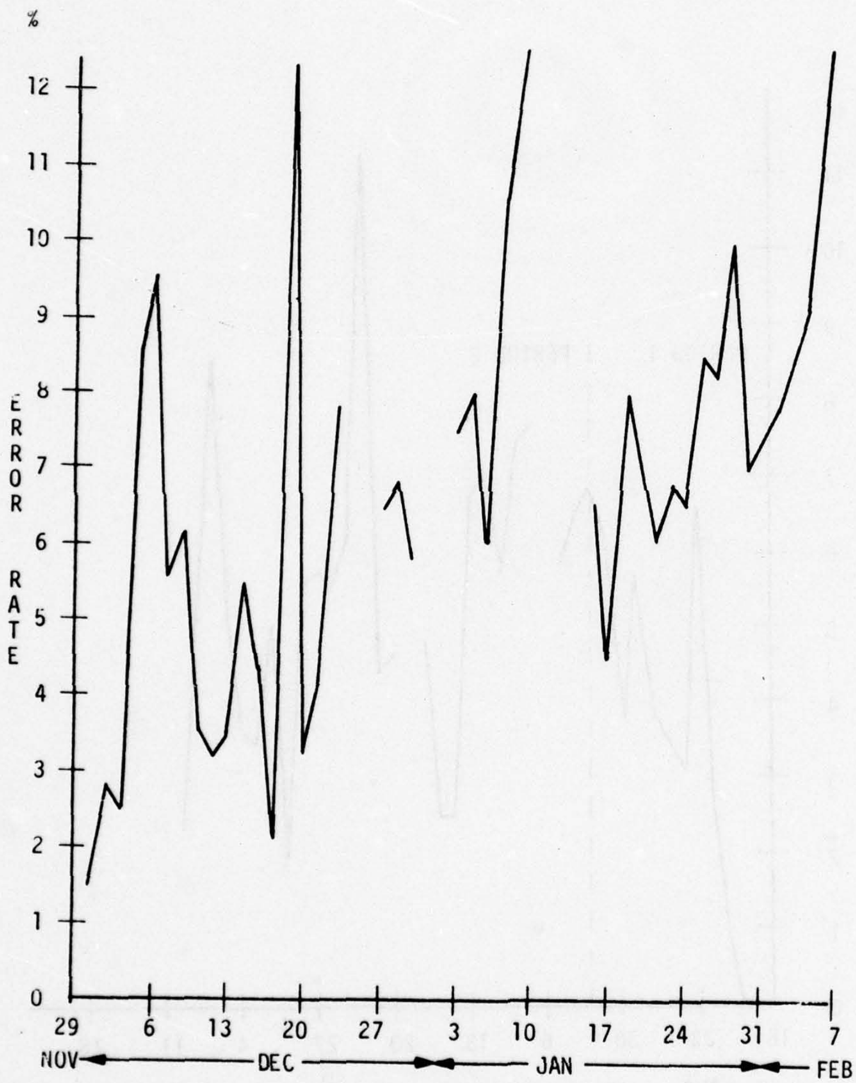


Figure 10. Daily Type I Error Rate - Field Test

Table III

Type I Error Rates

	PHASE II PERIOD 2	FIELD TEST
ATTEMPTS	3639	7152
TYPE I ERRORS	193	468
TYPE I ERROR RATE	5.30%	6.54%

The Type I data shows that the Field Test had a higher error rate than Phase II and submitting the data to an F ratio test* shows the error rates to be significantly different, i.e.,

$$T = \frac{3639}{7152} \times \frac{469}{194} = 1.23$$

$$F_{\alpha=.9}\ (388,938) \approx 1.11$$

In addition to a higher Type I error, the Field Test also shows a definite increasing error rate as a function of time which is not supported by Phase II (compare Figures 9 and 10). It is felt that the reason for these differences is because in Phase II there were many reenrollments and the test was 4-1/2 weeks shorter with 50% fewer verification attempts compared to the Field Test.

---

*If the parameter F is less than the test variable T than at the 90% confidence level, the two test groups have significantly different error rates. If F is greater than T then at the 90% confidence level, it cannot be said that the two test groups have significantly different error rates. See Appendix A.

6.1.2.1 <u>Confidence Limits of the Type I Error Estimate*</u>. The upper bound on the Type I error at 90% confidence is given by the Chi-squared test:

$$\text{Errors}_{0.9} = [\chi^2 \ (2e+2)]/2$$

where e = observed errors.

For values of (2e+2) > 100 that are not included in the table of Appendix B, an approximation is used where the value of

$$\chi^2_{.9} \ (2e+2) \approx 1/2 \ [1.286 + \sqrt{2\nu - 1}\ ]^2$$

where ν = (2e+2)

Using the approximation, a 90% confidence level on the upper bound of the Type I error rate may be determined for both test programs.

Phase II, Period 2:

$$\text{Errors}_{0.9} \approx 1/4 \ [1.286 + \sqrt{2(388)-1}\ ]^2 \approx 212$$

$$\text{Upper Bound Error Rate} \approx \frac{212}{3639} = 5.83\%$$

*See Appendix B.

Field Test:

$$\text{Errors}_{0.9} \approx 1/4\ [1.286 + \sqrt{2(938)-1}\ ]^2 = 497$$

$$\text{Upper Bound Error Rate} \approx \frac{497}{7152} = 6.95\%$$

Thus, the best estimate of the overall Type I error rate is 5.30% and 6.54% for Phase II and the Field Test, respectively, with a 90% confidence level on the upper bound of 5.83% for Phase II and 6.95% for the Field Test. This data is presented in Table IV.

Table IV

Confidence Limits for Type I Error Rates

	Phase II	Field Test
Total Attemtps	3639	7152
Total Errors	193	468
Type I Error Rate	5.30%	6.54%
Upper Bound with 90% Confidence	5.83%	6.95%

6.1.2.2 Weekly Type I Error Rates and Possible Causes. As people used the AFV system, the operators observed the verification process. When a person had a type I error, a Type I error message was sent to a teleprinter and the operator noted any circumstances that he had observed that may have contributed to the failure, e.g., used wrong finger, improper use of terminal, etc. In addition, the operator questioned the person that failed for possible reasons, e.g., cold hands, scarred finger, etc. Tables V and VI for Phase II and the Field Test, respectively, list the weekly number of verifications and

Table V

Phase II - Weekly Type I Error Data

Week Ending	Total Verifications	Total Failures	Type I Error %	Injured Finger	Wrong Finger	Consistent Failures	Other
Aug 20	258	10	3.88	0	0	3	7
Aug 27	635	28	4.41	0	0	1	27
Sept 3	901	53	5.88	3	1	11	38
Sept 10	859	60	6.98	4	0	5	51
Sept 17	673	27	4.01	2	0	2	23
Sept 24	873	42	4.81	0	0	10	32
Oct 1	939	40	4.26	0	0	7	33
Totals	5138	260	5.06	9	1	39	211
(From Table III)	(5039)	(261)	(5.18)				

Table VI

Field Test - Weekly Type I Error Data

Week Ending	Total Verifi-cations	Total Failures	Type I Error %	Cold Hands	Wrong ID	Wrong Finger	Difficul-ty in Using Terminal	Consis-sistent Failures	Other
Dec 3	603	20	3.32	3	0	2	1	14	0
Dec 10	820	46	5.61	6	0	3	1	12	24
Dec 17	950	42	4.42	3	2	1	2	10	24
Dec 24	375	19	5.07	4	1	1	2	0	11
Dec 31	334	22	6.59	7	1	0	3	0	11
Jan 7	1016	87	8.56	4	4	1	2	16	60
Jan 14	91	6	6.59	4	0	0	0	0	2
Jan 21	978	65	6.65	7	0	1	0	29	28
Jan 28	1116	89	7.97	1	4	6	0	19	59
Feb 4	952	84	8.82	1	1	0	0	20	62
TOTALS	7235	480	6.63	40	13	15	11	120	281
From Table III	(7152)	(468)	(6.54)						

failures, the reasons given or observed for failing, and their frequency of occurrence. It should be noted that not all those that failed were questioned or observed, and data is not available as to how many successful verifications occurred for people with cold hands, scarred fingers, etc. A discussion of each table follows:

TABLE V. In comparing the total verifications and failures between Table III and Table V, there is a discrepancy. This discrepancy arises because the data for Table III was taken from computer statistics. (The computerized data of Table III is included in parentheses in Table V for comparison.) The data for Table V was taken from log sheets and daily statistics and it is possible that slight errors were made when the data was manually recorded and collated. It is felt, however, that the differences are slight and therefore of no consequence.

<u>Injured Finger</u>. Of the 260 Type I errors, nine people responded that they may have failed because of injuries to fingers, e.g., cuts and scratches, or because of roughened hands acquired while gardening, doing cement work, etc.

<u>Wrong Finger</u>. Test participants were observed to use wrong fingers for both their Primary and Alternate finger. In Phase II this resulted in a single Type I error. However, there were 37 known cases where individuals were required to use their Alternate finger because they used a wrong finger for their Primary. As will be discussed in 6.2.6, the Alternate finger has a lower probability of passing verification than the Primary finger; therefore, the number of Type I errors attributable to persons using a wrong finger are higher than the single incident of Table V indicates. For the Field Test (Table VI), 15 incidences of people using a wrong finger for both their Primary and Alternate were noted. Occurrences of people using a wrong finger for just their Primary were not observed in the Field Test.

The total contribution of Type I errors due to wrong finger was difficult to identify, and it is felt that there were many undetected errors of this nature.

Consistent Failures. Three people with the highest Type I error rate were selected and the number of times per week that they caused a Type I error was analyzed. Choosing only the three highest was strictly arbitrary. They generated 15% and 25% of the Type I errors in Phase II and the Field Test respectively. These people, as did others, failed consistently even after several reenrollments. A photograph of a typical fingerprint image belonging to a person who failed often is shown in Figure 11. The fingerprint image is much more complex than the normal fingerprint image shown in Figure 12.

Other. In this column are the number of people who failed and were either not asked or when they were asked, no reason could be given for the failure. Included in this group are people who may have had damaged fingers, or used wrong fingers.

TABLE VI. Again, in comparing the total verification and failure data of Table III and Table VI, there are slight differences due mainly from the two ways in which the data was accumulated (manually and computer). The differences are slight, however, and have no bearing on the outcome or conclusions. In addition to the failure categories discussed for Phase II, there were several additional categories for the Field Test.

Cold Hands. There had been some concern prior to the Field Test that people with cold hands would be responsible for higher Type I error rates. During the Field Test, therefore, when people failed they were asked if their hands were cold. This was purely

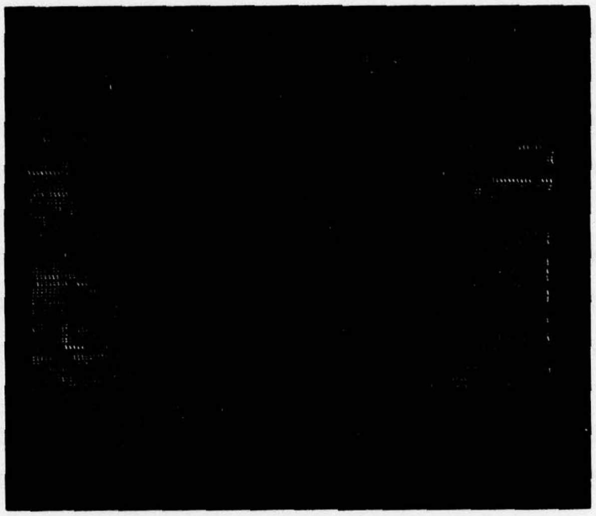


Figure 11. Fingerprint Image of a Person Who Failed Consistently.

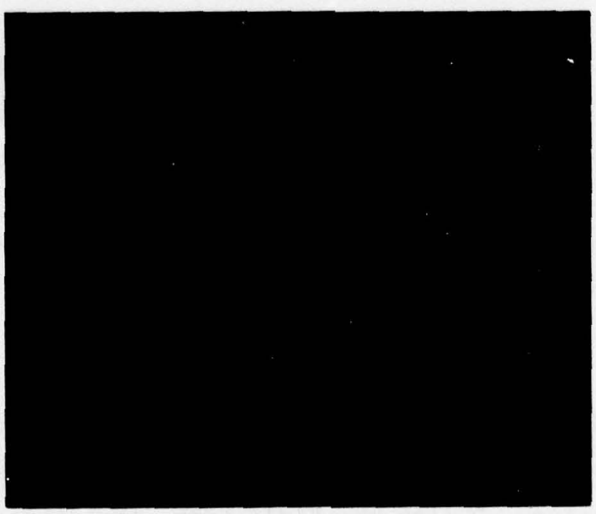


Figure 12. Fingerprint Image of a Person Who Passed Consistently.

a subjective evaluation, but of the 480 failees 40 gave a positive response. The number of people with cold hands that passed the verification could not be determined. This number most likely was quite high because of the time of year in which the Field Test was conducted and also because when entering the WSA, the test participants went immediately to the AFV terminal without having a chance to warm up. Cold hands was not a factor in Phase II because people worked indoors and the time of year was late summer and early fall.

The Field Test did experience higher Type I error rates than Phase II and most of the test participants in the Field Test had cold hands when verifying, which tends to support the hypothesis that people with cold hands will be responsible for a higher Type I error rate.

<u>Wrong ID Number</u>. Of the 480 Type I errors, 13 were due to instances of test participants attempting to verify on another person's reference file. This happened when a person wanting to verify made the mistake of entering another person's ID number.

It is hypothesized that the reason this occurred in the Field Test and not in Phase II was because of the difference in the distribution of the ID numbers. In the Field Test, the ID number of the 270 enrollees were closely distributed between 0005 and 0469. In Phase II, the ID numbers of the 200 enrollees were much more widely dispersed between 0052 and 9500. Intuitively one can see that a Field Test participant when making a mistake had a greater chance of entering another's ID than the Phase II participant.

Type I errors due to wrong ID was not unique to the AFV system in the Field Test; the other two systems experienced similar results.

Difficulty in Using Terminal. There were a few people who, due to the structure of their hand and/or fingers, had a difficult time in positioning their finger in the finger scanner. The number of people who had difficulty and also caused a Type I error are listed. Since no errors of this type were evident in the last four weeks of test, it would appear that the individuals were able to adjust to their particular problem.

6.2 OBJECTIVE 24 - TYPE I ERROR ANALYSIS IN NON-REAL TIME

6.2.1 Statement of Objective

To determine how various parameters affect the Type I errors and system performance.

6.2.2 Type I Error Rate vs Time of Day

The Type I errors as a function of the morning (AM) and afternoon (PM) are listed in Table VII for both males and females.

In all instances the error rate in the morning equals or exceeds the error rate in the afternoon. To test if the error rates between the morning and afternoon groups are significantly different, the F-ratio test is applied.

Phase II, Period 2, AM vs PM: $T = \frac{1601}{2038} \times \frac{130}{65} = 1.57$

$$F_{\alpha=.9}(130,260) \approx 1.20$$

Table VII

Type I Error Rate vs Time of Day

	Phase II						Field Test	
	Period 1		Period 2		Total			
	AM	PM	AM	PM	AM	PM	AM	PM
Males								
Attempts	533	600	1666	1284	2199	1884	4170	2989
Type I Error	31	18	91	44	122	62	278	158
Error Rate (%)	5.82	3.00	5.46	3.43	5.55	3.29	6.67	5.29
Female								
Attempts	127	140	372	317	499	457	274	209
Type I Error	9	10	38	20	47	30	21	11
Error Rate (%)	7.09	7.14	10.22	6.31	9.42	6.56	7.66	5.26
Totals (M+F)								
Attempts	660	740	2038	1601	2698	2341	3954	3198
Type I Error	40	28	129	64	169	92	299	169
Error Rate (%)	6.06	3.78	6.33	4.00	6.26	3.93	7.56	5.28

Field Test, AM vs PM.

$$T = \frac{3198}{3954} \times \frac{300}{170} = 1.43$$

$$F_{\alpha=.9}(340,600) \approx 1.11$$

With the results of the F ratio test, one can say with 90% confidence that the error rates between morning and afternoon are significantly different for both test programs. It has been suggested that this occurs because the hand is relatively dry and oil free in the morning.

6.2.3 Type I Error Rate vs Male and Female

The total error rates experienced by males and females for Phase II and the Field Test are presented in Table VIII.

Table VIII

Type I Error Rate vs Male and Female

	Phase II, Period 2		Field Test	
	MALE	FEMALE	MALE	FEMALE
Attempts	2950	689	6669	483
Type I Error	135	58	436	32
Error Rate	4.58%	8.42%	6.54%	6.63%

Though the results of the Field Test indicate a slightly higher Type I error rate for the females, the results are less significant statistically because of the small (6) female population using the system.

To determine if the error rates between males and females are significantly different in Phase II, the F-ratio test is applied.

Phase II, Male-Female:

$$T = \frac{2950}{689} \times \frac{59}{136} = 1.86$$

$$F_{\alpha=.9}(272,118) \approx 1.20$$

For Phase II, the F ratio test shows with 90% confidence that the error rates between males and females are significantly different.

An F-ratio test for the Field Test would not be a good indicator because of the low female population.

It is suggested that the reason females have a higher Type I error rate is because a female's fingerprint when compared to a male's has a finer ridge structure which may approach or exceed the resolution of the finger scanner.

6.2.4 Type I Error Rate vs Age Distribution

Table IX lists the Type I error rates as a function of a persons age for both test programs.

For Phase II, the verification attempts were predominantly made by people between the ages of 26 to 55 while for the Field Test the predominant ages ranged from 18 to 45.

Table IX

Type I Error vs Age

	Age Distributions				
	≤ 25	26-35	36-45	46-55	> 55
Phase II					
Attempts	435	717	1105	1111	271
Type I Errors	13	35	40	78	27
Error Rate (%)	2.99	4.88	3.62	7.02	9.96
Field Test					
Attempts	4181	1711	1160	96	4
Type I Errors	335	71	57	5	0
Error Rate (%)	8.01	4.15	4.91	5.21	0

One might anticipate that because of dryer skin, and/or more complex fingerprints due to scars, etc., that the Type I error rate would increase as a function of age. Phase II shows this increasing trend but in the Field Test the younger population ($\leq$25) experienced the highest error rate and there was insufficient data in the older categories to support the hypothesis that older people would be responsible for a higher Type I error rate.

6.2.5 Type I Error Rate vs Day of the Week

The Type I error rate as a function of the day of the week is shown in Tables X and XI for the entire Phase II and Field Test, respectively. For Phase II, the error rates appear variable as a function of the day of the week with no specific trend. For the Field Test, the overall error rate was lowest on Monday with a somewhat stable error rate for the middle of the week with a sharp increase on Friday. No reason could be found for the sharp increase in error rate on Friday.

The data for females is limited for both tests but is shown for comparison.

6.2.6 Number of Verification Attempts vs Scan Number for Decision

In a six scan decision strategy each attempt is given six scans in which to pass: three on the Primary finger and three on the Alternate finger. The attempt is considered a success on the first scan that passes the decision criteria. An attempt generates a Type I error when all six scans fail to pass. The total number of verification attempts that passed in each of the six scans for both test

Table X

Type 1 Error vs Day of Week - Phase II
(Periods 1 and 2)

	Day of Week				
	Mon	Tues	Wed	Thurs	Fri
Male					
Attempts	750	886	972	896	577
Type 1 Error	42	38	35	43	26
Error Rate (%)	5.6	4.3	3.6	4.8	4.5
Female					
Attempts	168	198	236	203	151
Type 1 Error	14	14	14	24	11
Error Rate (%)	8.3	7.1	5.9	11.8	7.3
Total					
Attempts	919	1084	1211	1090	735
Type 1 Error	56	52	59	67	37
Error Rate (%)	6.1	4.8	4.9	6.14	5.0

Table XI

Type 1 Error vs Day of Week - Field Test

	Day of Week				
	Mon	Tues	Wed	Thur	Fri
Male					
Attempts	1613	1508	1513	1114	921
Type 1 Error	90	100	99	69	78
Error Rate (%)	5.6	6.6	6.5	6.2	8.5
Female					
Attempts	105	97	113	99	72
Type 1 Error	12	7	5	5	3
Error Rate (%)	11.4	7.2	4.4	5.1	4.2
Total					
Attempts	1718	1605	1626	1213	993
Type 1 Error	92	107	104	74	81
Error Rate (%)	5.4	6.7	6.4	6.1	8.2

phases is presented in Table XII. Both tests had similar results with about 86% of the total attempts passing on the Primary finger and about 8% of the total attempts passing on the Alternate finger.

Of those attempts requiring the Alternate finger, 60% and 54% passed the decision criteria for Phase II and the Field Test respecitvely.

The data from Table XII can be plotted to show the probability of passing verification for each scan of a six scan decision strategy. The probability data is obtained for each scan by subtracting from the total number of attempts all previous successful scan attempts and dividing the remaining number into the number of successful attempts for the scan of interest.

The probability of passing verification for each of six scans is plotted in Figure 13. The plot is representative of both test programs and shows that the Type I error rate is dependent on scan number and that an individual requiring the use of the Alternate finger has a lower probability of passing verification.

6.2.7 Type I Error Rate Vs Primary and Secondary Matcher

As discussed in 4.2, one of two algorithms is assigned to an individual for matching fingerprint data with data in the reference file. The algorithm assigned most frequently is the Primary matcher with relatively few individuals requiring the Secondary matcher. The number of individuals assigned the Primary and Secondary matcher and the associated Type I errors generated is detailed in Table XIII for a three scan decision strategy. A three scan decision strategy is used so that the error rates for the Primary and Alternate finger

Table XII

Verification Attempts Passed vs Scan Number

Scan		Attempts Passed		% of Total Passing	
		Phase II	Field Test	Phase II	Field Test
PRIMARY	1	2485	4789	68.28	66.96
	2	507	998	13.93	13.95
	3	156	361	4.29	5.04
ALTERNATE	4	184	342	5.05	4.78
	5	69	119	1.90	1.66
	6	45	75	1.24	1.04
Failed		193	468	5.30	6.54
Totals		3639	7152		

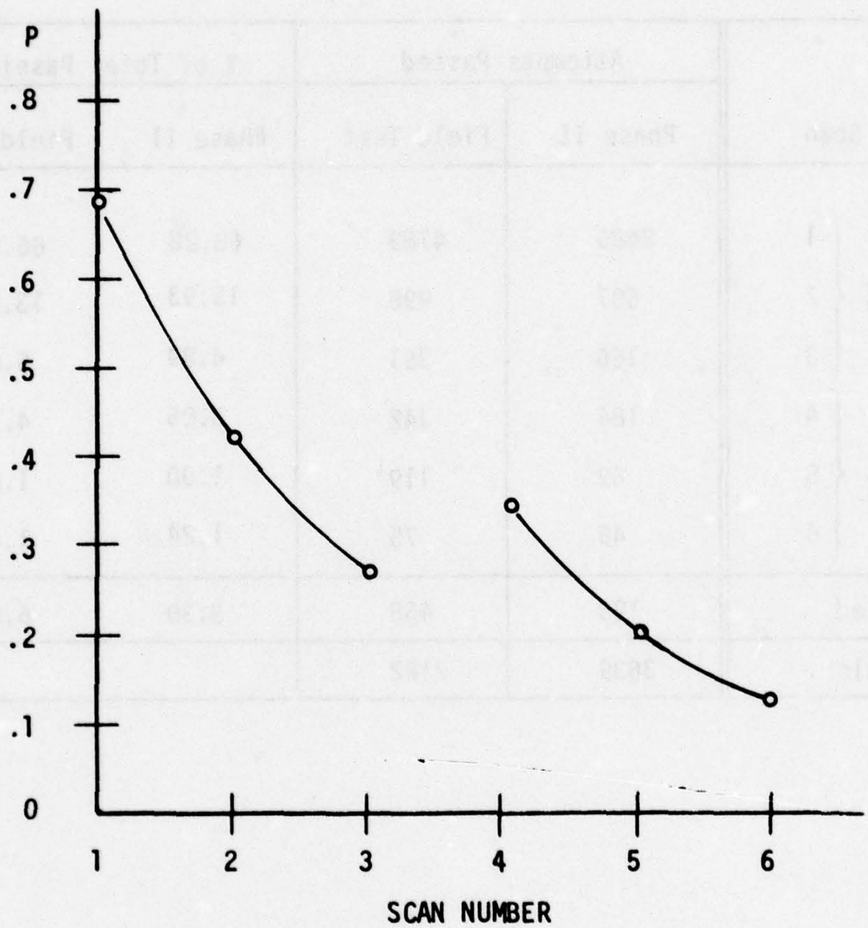


Figure 13. Probability of Verification vs Scan Number.

Table XIII

Type I Errors as a Function of the Primary and Secondary Matcher
(Three-Scan Decision Strategy)

	Phase II Period 2			Field Test		
	Attempts	Errors	Rate %	Attempts	Errors	Rate %
Primary Matcher						
Primary Finger	3,479	423	12.16	6,757	808	11.96
Alternate Finger	945	247	26.14	1,236	347	28.07
Totals	4,424	670	15.14	7,993	1,155	14.45
Secondary Matcher						
Primary Finger	160	68	42.50	395	196	49.62
Alternate Finger	74	22	29.73	321	221	68.85
Totals	234	90	38.46	716	417	58.24

can be determined separately. In a three scan decision strategy, A Type I error is generated when an attempted verification fails on only three scans as compared to six scans for a six scan decision strategy. The result of a three scan strategy, because there are fewer chances to pass, is a higher Type I error rate. The three scan decision strategy Type I data was generated by computer analysis of Phase II and the Field Test data.

In Phase II there were 22 individuals or 11.17% of the population that were assigned the Secondary matcher for either one or both of their Primary and Alternate finger(s). In the Field Test, there were 19 or 9.13% of the population that were assigned the Secondary matcher. In both Phase II and the Field Test, for a three scan strategy, individuals assigned the Secondary matcher were responsible for a considerably higher Type I error rate. To test if the Type I error rates between the Primary and Secondary matcher are significantly different, the F-ratio test is applied.

Phase II:

$$T = \frac{4424}{234} \times \frac{91}{671} = 2.56$$

$$F_{\alpha=.9}(1342, 182) \approx 1.15$$

Field Test:

$$T = \frac{7993}{716} \times \frac{418}{1156} = 4.04$$

$$F_{\alpha=.9}(2312, 836) \approx 1.07$$

Thus one can say with 90% confidence that the Primary and Secondary matchers have significantly different Type I error rates.

6.3 OBJECTIVE 25 - INDEPENDENCE OF TYPE I SCORES

6.3.1 Statement of Objective

To determine the independence of Type I errors for repeated uses of the system by individual enrollees as well as when compared against other enrollees.

6.3.2 Distribution of Type I Errors

From the real time data, the number of Type I errors and the frequency at which test participants made the errors were obtained. Tables XIV and XV present this data for both Phase II and the Field Test respectively.

Column 1 lists in consecutive order the number of Type I errors and column 2 lists the number of people responsible for that number of Type I errors. The total number of errors made by each group is listed in column 3 and is obtained by multiplying the numbers across columns 1 and 2; e.g., from Table XIV, 28 people out of the total participating population in Phase II each had a single Type I error for a total of 28 errors. The group of people that had no Type I errors include only those that verified after being enrolled. Those that were enrolled and never verified are not included.

From these tables, the data for plotting a distribution of errors is obtained. Figure 14 shows this distribution with about 4.5% of

Table XIV
Distribution of Type I Errors - Phase II Period 2

(1) Number of Type I Errors	(2) Number of People Responsible	(3) = (1)x(2) Total Type I Errors
0	135	0
1	28	28
2	16	32
3	2	6
4	3	12
5	3	15
6	1	6
7	0	0
8	1	8
9	3	27
10	1	10
11	2	22
12	1	12
13	0	0
14	0	0
15	1	15
16	0	0
Totals	197	193

Table XV

Distribution of Type 1 Errors - Field Test

(1) Number of Type I Errors	(2) Number of People Responsible	(3) - (1)x (2) Total Type I Errors
0	133	0
1	27	27
2	11	22
3	5	15
4	7	28
5	1	5
6	3	18
7	5	35
8	3	24
9	2	18
10	1	10
11	0	0
12	0	0
13	1	13
14	2	28
15	1	15
16	1	16
18	1	18
22	1	22
25	1	25
31	1	31
40	1	40
58	1	58
Totals	208	468

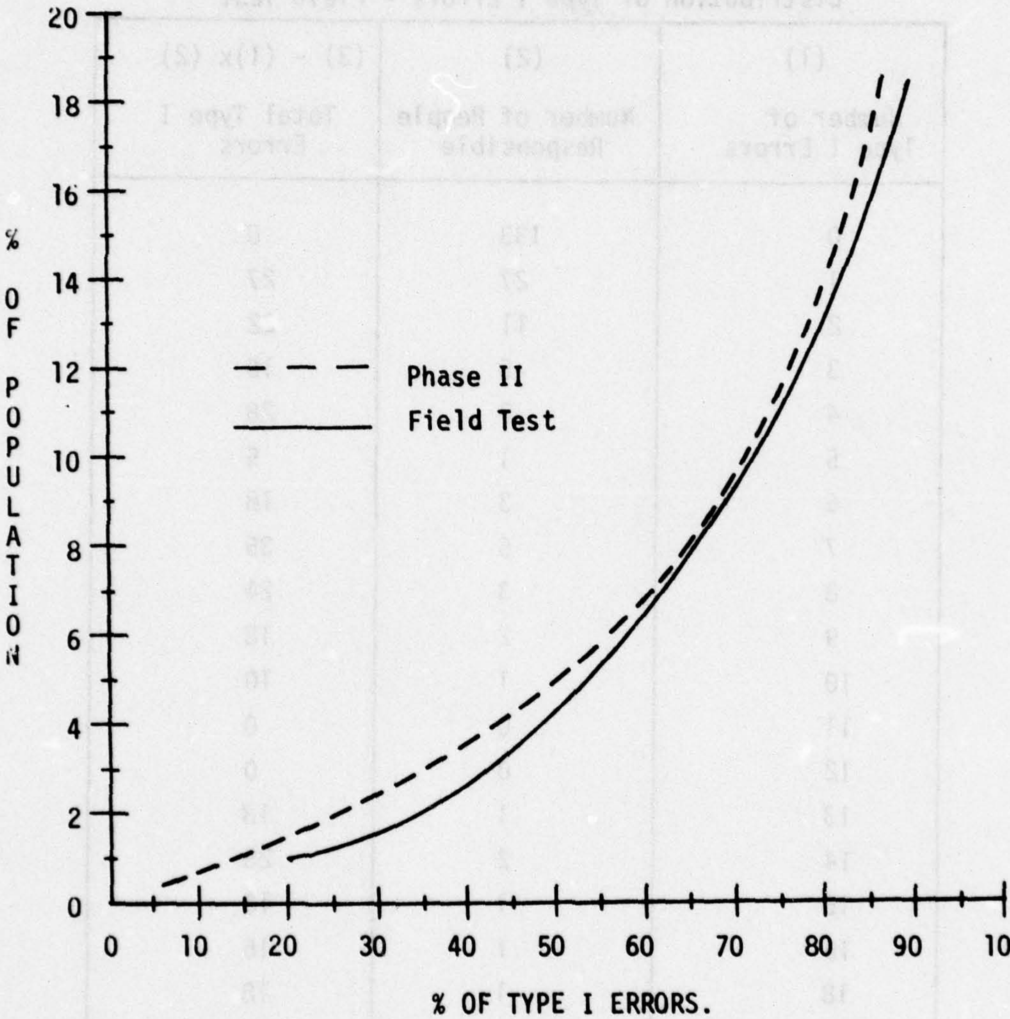


Figure 14. Distribution of Type I Errors.

the participating populations in Phase II and the Field Test responsible for about 50% of the Type I errors. The data shows that the Type I errors are heavily dependent on particular users and that the errors are not evenly distributed among the user population. Of the total test population, approximately 69% and 64% of the Phase II and Field Test respectively had no trouble passing the verification process. It was not uncommon, however, for individuals to have Type I error rates over 20% with several individuals approaching error rates of 60%.

6.4 OBJECTIVE 26 - TYPE II ERROR ANALYSIS IN REAL TIME

6.4.1 Statement of Objective

To estimate the Type II error (erroneously verifying identity when the representation is actually false) and to determine the confidence limits of the Type II error when imposters attempt verification.

6.4.2 Type II Error Analysis

During Phase II, various techniques were employed by individuals to generate Type II errors. The techniques used and the results are described in another document.

6.5 OBJECTIVE 27 - TYPE II ERROR ANALYSIS IN NON-REAL TIME

6.5.1 Statement of Objective

To estimate the Type II error rate (erroneously verifying identity when the representation is actually false) and to determine the confidence limits of the Type II error when matching within the enrolled population is performed.

6.5.2 Type II Error Rate

On every fourth verification attempt in Phase II and eighth verification attempt in the Field Test, imposter data was gathered for generating random Type II error rates. The imposter data consisted of 3 scans each of the Primary and Alternate finger in a form appropriate for both the Primary and Secondary matchers. This data was played back during the nighttime hours against all the reference files in an attempt to find a match between the imposter and all other enrollees.

The Type II error rates are based on a six scan decision strategy just as the Type I error rate. That is, each of the three scans of an imposter's Primary and Alternate fingers is compared respectively against Primary and Alternate finger scan data in each reference file, i.e., Primary vs Primary and Alternate vs Alternate. A Type II error occurred when any one of the six comparisons passed a decision threshold. The Type II error rate was determined using the formula:

$$\text{Type II Error Rate (\%)} = \frac{\text{Type II Errors}}{\text{Attempts}} \times 100$$

Figure 15 shows the daily Type II error rates for Phase II and the Field Test. There is no indication of an increasing trend in the Type II error rate as was evident in the Type I error rate for the Field Test (6.1.2).

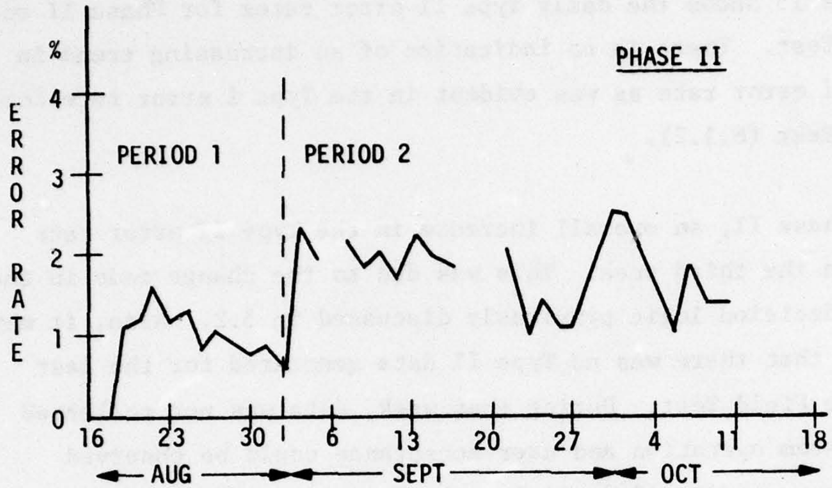
For Phase II, an overall increase in the Type II error rate occurred in the third week. This was due to the change made in the pass/fail decision logic previously discussed in 5.2. Also, it may be noticed that there was no Type II data generated for the last week of the Field Test. During that week, data was not collected so that system operation and user acceptance could be observed without the necessity of detaining the users to collect imposter data.

The Type II error rates for Phase II Period 2 and the Field Test are presented in Table XVI and XVII respectively. The tables divide the data into three comparisons: male vs male, male vs female, and female vs female.

In the male vs male comparison (i.e., male imposter data is compared against all male reference file data) the Type II error rate is slightly higher for the Field Test. To determine if the Type II error rate is significantly different between Phase II and the Field Test, an F-ratio test is performed.

$$T = \frac{48079}{60636} \times \frac{1405}{1050} = 1.06$$

$$F_{\alpha = .9}\ (2100,\ 2810) \approx 1.05$$



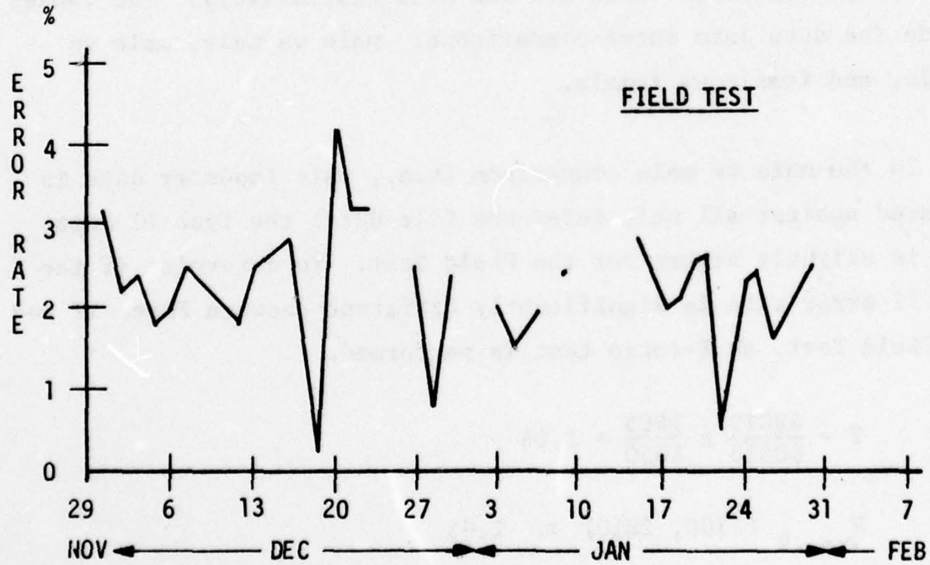


Figure 15. Daily Type II Error Rate

Table XVI

Total Type II Error Rates - Phase II

	Attempts	Errors	Error Rate
Male vs Male	48,079	1,049	2.18 %
Male vs Female	21,327	301	1.41 %
Female vs Female	2,279	75	3.29 %
Totals	71,685	1,425	1.99 %

Table XVII

Total Type II Error Rates - Field Test

	Attempts	Errors	Error Rate
Male vs Male	60,636	1,404	2.32 %
Male vs Female	5,647	91	1.61 %
Female vs Female	99	2	2.02 %
Totals	66,382	1,497	2.25 %

Thus, one can say with 90% confidence that the male groups in the Phase II and Field Test have significantly different Type II error rates. For comparison, the female data is presented in the table; however, it is felt that there were too few females involved in the Field Test to make a statistical comparison.

6.5.3 Confidence Limits of the Type II Error Estimate (Male Only)

Using the figures in Tables XVI and XVII, a 90% confidence level on the upper bound of the Type II error rate may be determined for both test programs.

Phase II, Period 2:

$$\text{Errors}_{0.9} \approx 1/4\ [1.286 + \sqrt{2(2100)-1}\ ]^2 = 1092$$

$$\text{Upper Bound Error Rate} \approx \frac{1092}{48{,}079} = 2.27\%$$

Field Test:

$$\text{Errors}_{0.9} \approx 1/4\ [1.286 + \sqrt{2(2810)-1}]^2 = 1454$$

$$\text{Upper Bound Error Rate} \approx \frac{1454}{60{,}636} = 2.40\%$$

Thus, the best estimate for the overall Type II error rate for males is 2.18% and 2.32% for Phase II and the Field Test respectively with a 90% confidence level on the upper bound of 2.27% for Phase II and 2.40% for the Field Test. This data is presented in Table XVIII.

Table XVIII

Confidence Limits for the Type II Error Rate (Male Only)

	Phase II	Field Test
Attempts	48,079	60,636
Errors	1,049	1,404
Error Rate	2.18%	2.32%
Upper bound with 90% confidence	2.27%	2.40%

6.5.4 Type II Error Rate Vs Primary and Secondary Matcher

As previously discussed in 6.2.7 approximately 11% and 9% of the participants in the Phase II and Field Test were assigned the Secondary matcher. Table XIX presents the data for comparing the performance of the Primary and Secondary matcher for Type II errors for a three scan decision strategy.

In both Phase II and the Field Test those individuals assigned the Secondary matcher were responsible for a lower Type II error rate. To test if the Type II error rates between the Primary and Secondary matcher are significantly different the F-ratio test is applied.

Phase II:

$$T = \frac{6620}{193,743} \times \frac{2196}{39} = 1.92$$

$$F_{\alpha=.9}(78,4392) \approx 1.21$$

Table XIX

Type II Error Rate vs Primary and Secondary Matcher
(Three Scan Decision Strategy)

	Phase II Period 2			Field Test		
	Attempts	Errors	Rate(%)	Attempts	Errors	Rate (%)
Primary Matcher						
Primary Finger	98,610	1,261	1.28	115,756	1,431	1.24
Alternate Finger	95,133	934	.98	114,352	1,340	1.17
Totals	193,743	2,195	1.13	230,108	2,771	1.20
Secondary Matcher						
Primary Finger	2,104	2	.09	5,714	11	.19
Alternate Finger	4,516	36	.80	7,154	18	.25
Totals	6,620	38	.57	12,868	29	.23

Field Test:

$$T = \frac{12868}{230{,}108} \times \frac{2772}{30} = 5.17$$

$$F_{\alpha=.9}(60{,}5544) \approx 1.25$$

Thus, one can say with 90% confidence that the Primary and Secondary matcher have significantly different Type II error rates.

6.5.5 Type II Error Rate vs Imposter Finger

In deriving the Type II errors for Phase II and the Field Test, all fingerprint data in the reference files are scanned using a six scan decision strategy in an attempt to find a match with the imposter fingerprint. Thus, if an imposter's Primary finger is the left index, that finger was matched to all Primary fingers in the reference file, i.e., right and left index, middle, ring, and little finger.

Data, derived from computer analysis, is presented in Table XX that shows the Type II error rates for the Primary and Alternate finger as a function of comparing an Imposter finger with the same Primary or Alternate finger of a victim (right index against right index, left middle against left middle, etc.). The data in Table XX are for a three scan decision strategy and only male-male and female-female comparisons are made. An approximation can be made for the Type II error rate that could be expected in a six scan decision strategy by summing the error rates for the Primary and Alternate finger. These totals are included in Table XX. In this analysis, the female-female comparisons had little effect on the overall error rates because of the low female population.

Table XX

Type II Error Rate vs Like Fingers of Impostor and Victim
(Three Scan Decision Strategy)

	Phase II Period 2			Field Test		
	Attempts	Errors	Rate %	Attempts	Errors	Rate %
Primary Finger						
Male vs Male	24,746	524	2.12	25,509	651	2.55
Female vs Female	1,257	23	1.83	41	1	2.44
Total	26,003	547	2.10	25,550	652	2.55
Alternate Finger						
Male vs Male	22,473	380	1.69	24,507	613	2.50
Female vs Female	1,014	24	2.37	35	0	0
Total	23,487	404	1.72	24,542	613	2.50
Approximate Type II Error rate to be expected for a six-scan strategy			3.82			5.05
Measured Type II Error rate (excluding male vs female) for six-scan strategy			2.23			2.32

The data shows (e.g., 3.82% vs 2.23% and 5.05% vs 2.32%) that an imposter can improve his chances of success by knowing which fingers have been assigned as the Primary and Alternate finger of his victim.

6.6 OBJECTIVE 28 - INDEPENDENCE OF TYPE II SCORES

6.6.1 Statement of Objective

To determine the independence of the Type II scores for repeated uses of the system by individual enrollees as well as when compared against other enrollees.

6.6.2 Distribution of Type II Errors

The distribution of Type II errors among the test populations is obtained in the same way as for the distribution of Type I errors described in 6.3.2. Tables XXI and Table XXII present the distribution data for the Phase II and Field Test respectively. Figure 16 plots the distribution of Type II errors for both test phases and shows that the distribution of Type II errors is not uniform across the test population with about 14% of the population responsible for 50% of the Type II errors. In addition, 12.5% of the Phase II population and 21.43% of the Field Test population had no Type II errors. The Type II errors are dependent on particular users and are not evenly distributed among the population. It was not unusual for individuals to be responsible for Type II error rates exceeding 10%.

Table XXI

Distribution of Type II Errors - Phase II Period 2

Number of Type II Errors	Number of People Responsible	Total Type II Errors
0	25	0
1	19	19
2	24	48
3	23	69
4	16	64
5	9	45
6	10	60
7	10	70
8	4	32
9	5	45
10	11	110
11	5	55
12	3	36
13	4	52
14	3	42
15	5	75
16	4	64
17	3	51
18	3	54
19	1	19
20	3	60
21	1	21
22	1	22
26	1	26
27	1	27
31	1	31
32	1	32
33	1	33
38	1	38
39	1	39
86	1	86
Totals	200	1425

Table XXII

Distribution of Type II Errors - Field Test

Number of Type II Errors	Number of People Responsible	Total Type II Errors
0	57	0
1	28	28
2	33	66
3	23	69
4	16	64
5	16	80
6	18	108
7	9	63
8	5	40
9	9	81
10	7	70
11	5	55
12	5	60
13	8	104
14	4	76
15	4	60
16	1	16
17	2	34
18	1	18
20	3	60
21	1	21
22	1	22
23	1	23
24	1	24
25	1	25
26	1	26
28	1	28
30	1	30
32	1	32
36	1	36
37	1	37
41	1	41
Totals	266	1497

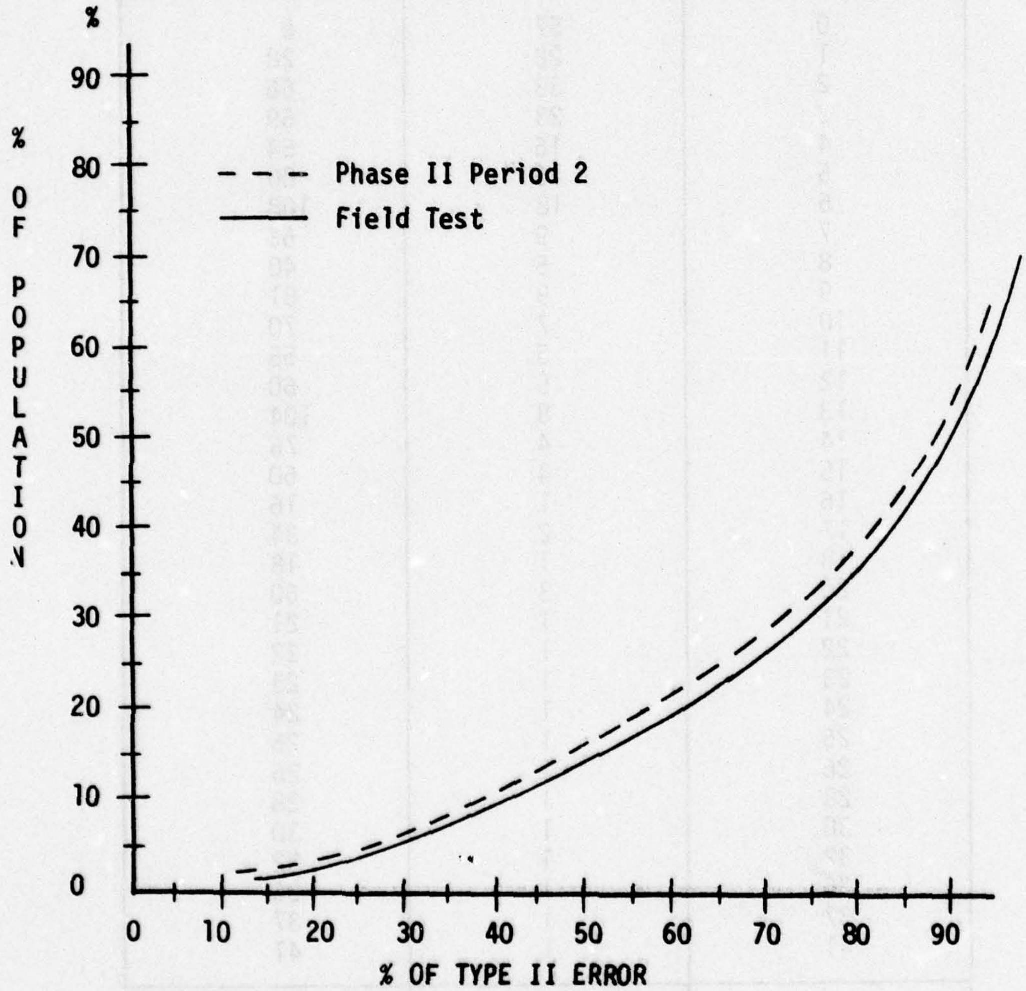


Figure 16. Distribution of Type II Errors.

6.7 OBJECTIVE 29 - SENSITIVITY ANALYSIS OF TYPE I AND TYPE II ERRORS TO THE DECISION THRESHOLD

6.7.1 Statement of Objective

To determine the sensitivity of Type I and Type II error variations to changing thresholds.

6.7.2 Sensitivity Analysis

For a Type I error to occur, the score obtained from matching an individual's fingerprint with the same individual's data in the reference file must be less than the pass/fail decision threshold. For a Type II error to occur, the score obtained from matching an impostor's fingerprint data with that of his victim's must be equal to or greater than the pass/fail decision threshold. There is a trade-off, therefore, between Type I and Type II errors and the level of the pass/fail decision threshold: as the threshold is raised, the system will experience more Type I errors at the expense of fewer Type II errors and vice versa.

To evaluate the sensitivity of Type I and Type II errors, the Field Test data was collated to show the percentage of errors occurring for the Primary matcher on the first, second and third scans as a function of the threshold level. Figure 17 shows the data plotted for each of the first three scans. Data for Phase II is not shown because it is similar to the Field Test data. When an individual's score falls below the decision threshold (score minus the threshold is negative), a Type I error is generated. When an impostor's score is equal to or above the decision threshold (score minus the threshold is zero or positive), a Type II error is generated. The Type I error

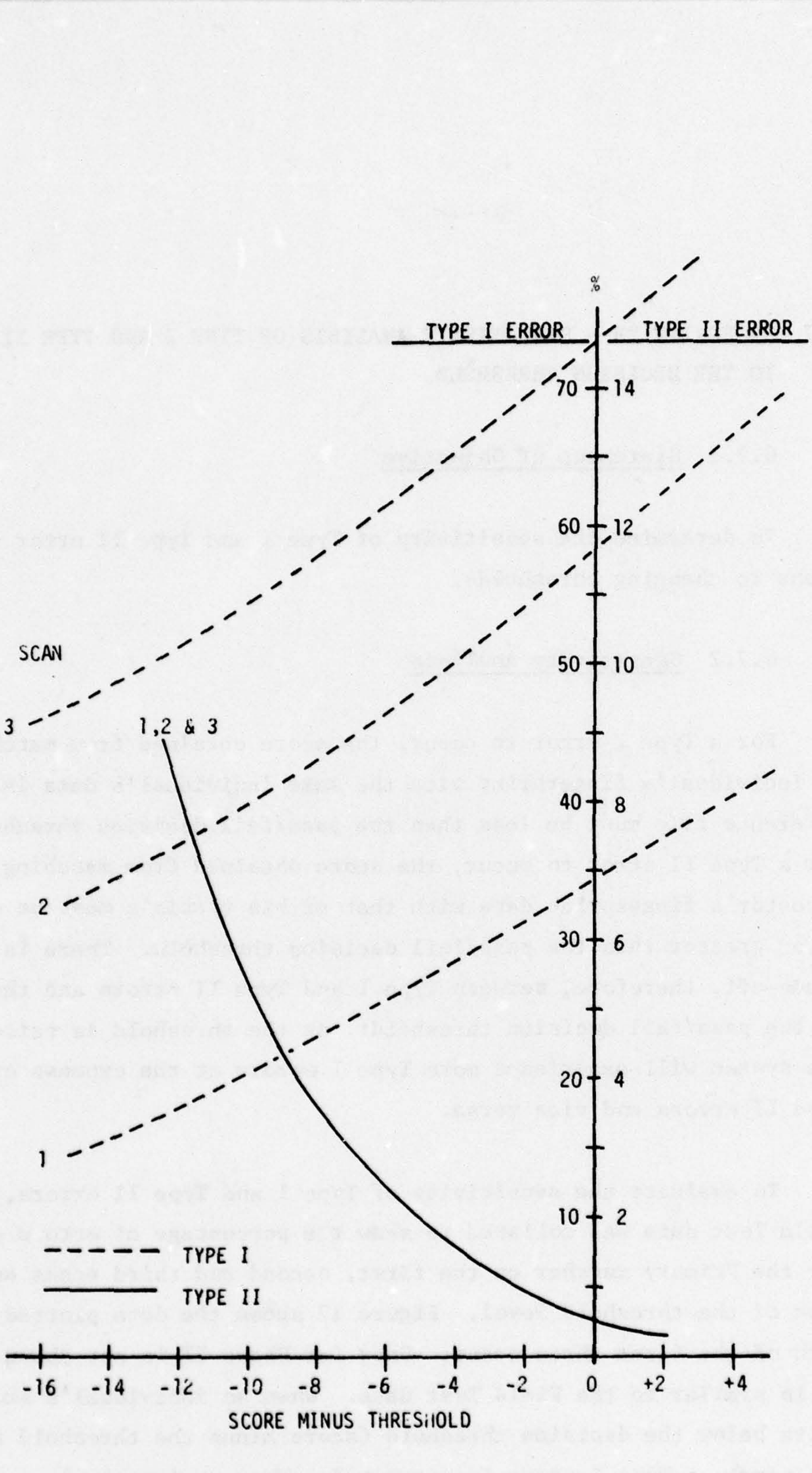


Figure 17. Type I and Type II Error vs Score Minus Threshold Field Test

rate is shown to be dependent on the scan number, e.g., from Figure 17 at a score minus threshold = 0 there is a 33% Type I error rate for scan 1, a 58% Type I error rate for scan 2 and a 74% Type I error rate for scan 3. The Type II error rate of Figure 17 is shown to be independent of scan number, e.g., the Type II error rate is approximately the same for the first, second, and third scans.

The data in Figure 17 can be configured to show the overall sensitivity of the Type I and Type II error rate to threshold for a three scan decision strategy. In a three-scan decision strategy, a Type I error is generated when an individual fails all three scans. Therefore, to obtain Type I error data for a three-scan decision strategy, the percentages in each of three scans as a function of score minus threshold are multiplied together; e.g., for the Field Test, at a score minus threshold = 0, 33% for scan 1 x 58% for scan 2 x 74% for scan 3 = 14.16% error rate for a three scan decision strategy.

A Type II error is generated in a three-scan decision strategy when an imposter matches a reference file on any one of the three scans. Therefore, to obtain a Type II error curve showing the performance of a three-scan decision strategy, the percentage of people in each scan as a function of score minus threshold is approximated by summing the errors in each scan, e.g., for the Field Test, a score minus threshold = 0, .6% for scan 1 + .6% for scan 2 + .6% for scan 3 = 1.8% error rate for a three-scan decision strategy.

The results of a three-scan decision strategy and the sensitivity of the error rates to threshold for the Primary matcher are shown in Figure 18 for the Field Test. The Phase II results would be similar. Ideally, it would have been preferred to show the same type of

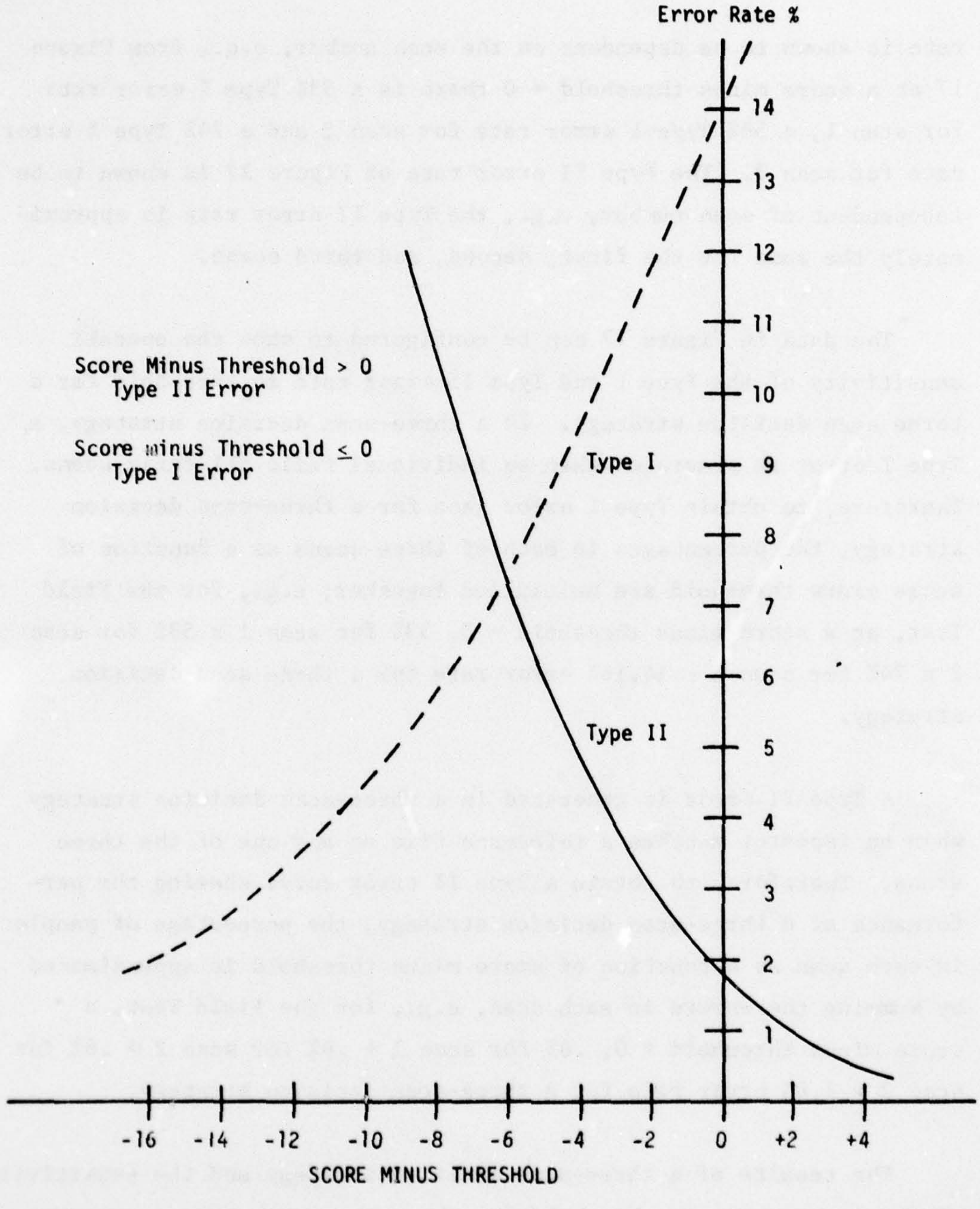


Figure 18. Sensitivity of Type I and Type II Errors to Threshold (Three Scan Strategy - Field Test)

analysis for the six-scan decision strategy used in Phase II and the Field Test and included the data for the Secondary matcher for both Primary and Alternate fingers, but data was not available to do this. Had data been available, an analysis similar to the one just discussed for the Primary matcher could have been performed for the Secondary matcher. The data could then have been combined to generate Type I and Type II error curves for a six-scan decision strategy employing both Primary and Secondary matchers. A complete analysis such as this would have shown the Type I error curve of Figure 18 intersecting the ordinate at 6.54% (Section 6.1.2) and the Type II error curve intersecting the ordinate at 2.32% (Section 6.5.2). It is evident that if one elected to lower the Type I error rate to better meet BISS requirements by lowering the threshold (i.e., in Figure 18 causing the Type I and Type II error curves to move to the right), the Type II error rate would increase rapidly and further exceed the BISS requirement of 2%.

6.8 OBJECTIVE 30 - VERIFICATION TIME ANALYSIS

6.8.1 Statement of Objective

To determine the average time required for verification and the variance about this time.

6.8.2 Verification Time

Table XXIII lists the parameters for calculating the time from the moment an ID number is accepted until a THANK YOU or TRY FINGER AGAIN message is displayed. The data does not include the time required to process imposter data nor does it include the time required for the users to enter their identification numbers.

Table XXIII

Parameters for Determining Verification Time

	PHASE II	FIELD TEST
Total No. of Scans (St)	9,019	13,089
Total Attempts (At)	5,039	7,152
Total Time	34,951 sec.	52,100 sec.
Avg. Time/Scan (t)	3.88 sec.	3.98 sec.
Standard Deviation	0.9 sec.	1.06 sec.

To obtain the average time for a pass/fail decision to be made, the average time per scan is multiplied by the number of scans required for a decision. Also to be factored into the time element, is the time it takes for an user, after the first scan, to remove and reinsert his finger and press the ENTER button on the keypad. This time has been estimated to take 2 seconds. With the above information the following formula is used to determine the verification time:

$$V_T = 2(N-1) + (N \times t)$$

where

V_T = Verification Time

N = Average Number of scans per attempt = $\frac{St}{At}$

t = Average time per scan

For Phase II:

$$V_T = 2(1.79 - 1) + (1.79 \times 3.88) = 8.53 \text{ seconds}$$

and for the Field Test:

$$V_T = 2(1.83 - 1) + (1.83 \times 3.98) = 8.94 \text{ seconds.}$$

7.0 CONCLUSIONS

From the results reported herein on both Phase II and the Field Test, the following conclusions are made:

CONCLUSIONS AS SUPPORTED BY BOTH TEST PHASES	SUPPORTING PARAGRAPHS
1. The AFV Type I error rate:	
a. Does not meet the BISS requirement of 1%	6.1
b. Is significantly higher in the morning than in the afternoon	6.2.2
c. For the Primary finger, was much better (≈14% Phase II and Field Test) than it was for the Alternate finger (≈40% Phase II and ≈46% Field Test).	6.2.6
d. Is dependent on scan number	6.2.6
e. For the Primary matcher was much better (≈15% Phase II and Field Test) than it was for the Secondary matcher (≈38% Phase II and ≈58% Field Test)	6.2.7
f. Is not evenly distributed among the user population with approximately 5% of all users generating 50% of the Type I errors and with approximately 22% of the users having a Type I error rate exceeding 5%.	6.3
2. The AFV Type II error rate:	
a. Does not meet the BISS requirement of 2%	6.5.2
b. Does not show an increase with time since enrollment	6.5.2
c. For the Primary matcher was significantly higher (1.13% for Phase II and 1.20% for the Field Test) than it was for the Secondary matcher (.57% for Phase II and .23% for the Field Test).	6.5.4

CONCLUSIONS AS SUPPORTED BY BOTH TEST PHASES	SUPPORTING PARAGRAPHS
d. Will be higher for an imposter who knows the Primary and Alternate finger of his victim.	6.5.5
e. Is not evenly distributed among the user population with approximately 14% of the population generating 50% of the Type II errors.	6.6
f. Is independent of scan number	6.7
3. The average verification time, excluding ID number entry and finger repositioning, is approximately 8.7 seconds.	6.8

In addition, the Field Test data (6.1.2) shows that the Type I error rate increases with time. Initially, the Type I error rate was about 4% but by the end of the Field Test it was close to 8%. Increasing error rates with time since enrollment were not observed to rise as rapidly in Phase II perhaps because (1) there was a parameter change part way into the test, (2) there were more re-enrollments performed in Phase II than in the Field Test, and (3) the Field Test was 4-1/2 weeks longer in duration than Phase II Period 2.

The Phase II test data suggests that the Type I and Type II error rates are higher for females than for males. Because of the limited number of females in the Field Test, a comparison could not be made.

It has been hypothesized that persons with cold hands will experience a higher Type I error rate. Because most test participants in the Field Test had cold hands and the Field Test did have higher Type I error rates, the hypothesis has not been disproved.

Finally, the personnel statistics of age, height, birth place, and area of primary education did not exhibit any consistent influence on the error rates.

APPENDIX A

DETERMINING IF THE ERROR RATES FROM TWO GROUPS ARE SIGNIFICANTLY DIFFERENT

In BISS data analysis it is often desired to determine if the error rates from two groups are significantly different. The BISS data provides only samples from different groups. The F test provides a test to determine statistically if the underlying populations of the different groups do not have the same error rates.

The binominal distribution

$$P_N(k) = \frac{N!}{k!(N - k)!} p^k (1 - p)^{N-k} \qquad (1)$$

is the probability of having k errors in N trials where the probability of error is p and that of no error is 1 - p. Each trial is taken to be independent. In the limiting case where

p is small ($p \ll 1$)

N is large ($N \gg 1$) and

$$N \gg k, \qquad (2)$$

the binomial distribution may be approximated by the Poisson distribution as[1]

$$\frac{N!}{k!(n-k)!} p^k (1-p)^{N-k} \approx \frac{(Np)^k e^{-Np}}{k!} \tag{3}$$

From the relationship

$$\Gamma(x+1) = x!$$

equation (3) may be rewritten as

$$P_k(Np) = \frac{(Np)^{(k+1)-1} e^{-Np}}{\Gamma(k+1)} \tag{4}$$

The continuous variable Np is a Gamma variate with parameter k + 1 and its distribution, equation (4), is a Gamma distribution.[2]

The probability that the variable Np is in the interval dNp is

$$dP = \frac{(Np)^{(k+1)-1} e^{-Np}}{(k+1)} dNp \tag{5}$$

Substituting

$$k + 1 = \frac{n}{2} \tag{6}$$

into equation (5) yields:

$$dP = \frac{2^{\frac{1}{2}(n-2)} (Np)^{\frac{1}{2}(n-2)} e^{-Np}}{2^{\frac{1}{2}(n-2)} \Gamma(n/2)} d(2Np/2)$$

$$= \frac{(2Np)^{\frac{1}{2}(n-2)} e^{-Np}}{2^{n/2} \Gamma(n/2)} d(2Np) \qquad (7)$$

The probability density function

$$P_C(2Np) = \frac{(2Np)^{\frac{1}{2}(n-2)} e^{-Np}}{2^{n/2} \Gamma(n/2)}$$

is the chi-square distribution for the variate 2Np with n degrees of freedom.

Consider that $2N_1p_1$ and $2N_2p_2$ are independent, and that each is chi-square distributed with n_1 and n_2 degrees of freedom, respectively. Then

$$T(n_1,n_2) = \frac{2N_1p_1 \;/\; n_1}{2N_2p_2 \;/\; n_2} \qquad (8)$$

is the F variate and its distribution is the F distribution.

Replacing n with 2(k + 1) from equation (6), equation (8) becomes

$$T[2(k_1 + 1), 2(k_2 + 1)] = \frac{2N_1 p_1 (2k_2 + 2)}{2N_2 p_2 (2k_1 + 2)} = \frac{N_1 p_1 (k_2 + 1)}{N_2 p_2 (k_1 + 1)} \quad (9)$$

If the Error rates p_1 and p_2 of the underlying populations are hypothesized to be equal, then equation (9) becomes

$$T(2k_1 + 2, 2k_2 + 2) = \frac{N_1(k_2 + 1)}{N_2(k_1 + 1)}$$

This hypothesis can be tested by seeing how often it will happen that the F variate with parameters $2k_1 + 2$ and $2k_2 + 2$, $F(2k_1 + 2, 2k_2 + 2)$, will take on values $\geq (N_1 k_2 + N_1)/(N_2 k_1 + N_2)$.

There are tables (3) of the cumulative F distribution tabulating against a_1, a_2 the values of $F(a_1, a_2)$ such that

$$P = \int_0^{F_P} f[F(a_1, a_2)]\, dF = \int_0^{F_P} f[F(2k_1 + 2, 2k_2 + 2)]\, dF \quad (10)$$

where f(F) is the probability density function. From equation (10) it is meant that 100P % of the time the values of $F(2k_1 + 2, 2k_2 + 2)$ will be $\leq F_P$. Thus, if the observed T is less than or equal to F_P, then there is a 100P % "level of confidence" and $T = (N_1 k_2 + N_1)/(N_2 k_1 + N_2)$ is an F variate and therefore the underlying hypothesis of equal error rates cannot be rejected.

If P is chosen large enough, then in the case of $T > F_P$ it can be concluded that it is very unlikely that T is F distributed. The hypothesis of equal error rates can then be rejected. The inverse reasoning is not possible. If $T \leq F_p$, one cannot say that T is F distributed and the hypothesis is true, but only that the hypothesis is not inconsistent with the results.

For example, if N_i are the entry attempts of group i and k_i are the errors in group i, then for

$$N_1 = 1646 \qquad k_1 = 23$$

$$N_2 = 1027 \qquad k_2 = 19$$

$$T[2(23 + 1), 2(19 + 1)] = \frac{1646\ (19 + 1)}{1027\ (23 + 1)}$$

or

$$T(48, 40) = 1.336$$

From the F_p tables for P = .9, the value of F_p such that

$$.9 = \int_0^{F_P} f[F(48, 40)]\, dF$$

is

$$F_p = 1.49$$

For this example, $T \leq F_{p = .9}$ and therefore, at the 90% confidence level the hypothesis that the two groups have equal error rates cannot be rejected.

In this manner, the F test can be used to determine if the error rates of two different groups are significantly different in a statistical sense.

APPENDIX B

COMPUTING AN UPPER BOUND ON ERRORS

The basis for the formula to compute error bounds in identity verfication is described below. Assume that the number of verification errors for a given number of (trials) entry attempts is statistical equivalent to the number of failures of a piece of equipment for a given number of operating hours. Specifically, failures are replaced with errors and operation hours are replaced with entry attempts. Now, the upper bound on the number of errors at a given confidence level can be estimated for a fixed number of trials in the same way that the number of failures, at a given confidence level, can be estimated for a fixed number of operating hours.

The formula used to estimate the lower bound of the mean time between failures (MTBF) for a fixed time test is given by[4]

$$MTBF_c = 2T \Big/ [\chi^2_{1-c}\ (2f + 2)].$$

where

T is time period

c is confidence level, e.g., 90%

f is number of failures

2f + 2 number of degrees of freedom

The upper bound on failure rate is the inverse of the lower bound on MTBF. Thus

AD-A056 918 MITRE CORP BEDFORD MASS F/G 15/3
TEST RESULTS-ADVANCED DEVELOPMENT MODELS OF BISS IDENTITY VERIF--ETC(U)
JUL 78 P BENSON F19628-77-C-0001
UNCLASSIFIED MTR-3442-VOL-4 ESD-TR-78-150-VOL-4 NL

2 OF 2
AD
A056 918

END
DATE
FILMED
9-78
DDC

$$\text{Upper Bound Failure Rate} = 1/\text{MTBF}_{\text{lower}} = \chi^2\ (2f + 2)/2T$$

The upper bound on the number of failures divided by the time period T is the upper bound on the failure rate

$$\text{Failures}_c = \chi^2_{1-c}\ (2f + 2)/2.$$

Substituting errors, e for failures, f yields

$$\text{Errors}_c = \left[\chi^2_{1-c}\ (2e + 2)\right]/2.$$

An example of the use of this equation follows. On the first entry attempt using the speaker verification system there were 8 failures to verify an authorized individual out of 193 entry attempts leading to an e = 8. The upper bound on the error at 90% confidence is given by (see Reference 5 for chi-square distribution tables)

$$\text{Errors}_{.9} = \chi^2(18)/2 = 26/2 = 13.$$

We can say with 90% confidence that no more than 13 errors out of 193 trials will occur on the first trial.

The best estimate of the error rate is 4.14% and a 90% confidence level on the upper bound of the error rate is 4.14% × 13/8 = 6.73%.

LIST OF REFERENCES FOR APPENDICES

1. Meyer, Stuart L., "Data Analysis for Scientists and Engineers", John Wiley and Sons, Inc. 1975, pp. 202-212, 285-286.

2. Weatherburn, G.E., "A First Course in Mathematical Statistics," Cambridge Press, 1949, Chapters VIII and IX.

3. Brandt, S., "Statistical and Computational Methods in Data Analysis," North-Holland Publishing Co., 1970, pp. 210-213.

4. Epstein, B., "Estimation from Life Test Data," IRE Transactions on Reliability and Quality Control, April 1960.

5. Handbook of Tables for Probability and Statistics, Beyer, W.H., Editor, the Chemical Rubber Co., 2,310 Superior Avenue, Cleveland, Ohio, 44114, pp. 234 and 241.