CONF-950787--11
SAND95-1297C

ENTRY CONTROL TECHNOLOGY BIOMETRIC FIELD EVALUATIONS

J. R. Rodriguez
Sandia National Laboratories
Albuquerque, New Mexico USA

J. S. Ahrens
Sandia National Laboratories
Albuquerque, New Mexico USA

D. L. Lowe
Sandia National Laboratories
Albuquerque, New Mexico USA

ABSTRACT

Laboratory evaluations provide system engineers basic design information. Laboratory generated data is useful for initial comparisons of devices and in estimating the effectiveness of a device within a system. Though laboratory data is very useful, it only presents a part of the total picture. Since the performance of a device may change in transition from lab to field, system engineers also require field data. Throughout the years, Sandia National Laboratories (SNL) has performed various laboratory evaluations of entry control devices, including biometric identity verifiers. The reports which resulted from this testing have been very well received by the physical security community. This same community now requires equally informative field study data. To meet this need we have conducted a field study in an effort to develop the tools and methods which our customers can use to translate laboratory data into operational field performance.

The field testing described in this report was based on the Recognition Systems Inc.'s (RSI) model ID3D HandKey biometric verifier. This device was selected because it is referenced in DOE documents such as the Guide for Implementation of the DOE Standard Badge and is the de facto biometric standard for the DOE. The ID3D HandKey is currently being used at several DOE sites such as Hanford, Rocky Flats, Pantex, Savannah River, and Idaho Nuclear Engineering Laboratory. The ID3D HandKey was laboratory tested SNL. It performed very well during this test, exhibiting an equal error point of 0.2 percent.

The goals of the field test were to identify operational characteristics and design guidelines to help system engineers translate laboratory data into field performance. A secondary goal was to develop tools which could be used by others to evaluate system effectiveness or improve the performance of their systems. Operational characteristics were determined by installing a working system and studying its operation over a five month period. Throughout this test we developed tools which could be used by others to similarly gauge system effectiveness.

1.0 SYSTEM REQUIREMENTS

Building Description

Building 956 is located just south of Sandia's Technical Area I and is primarily used by Sandia's security force and employees of the technical security organization. There are over 250 persons authorized access into the building. The building houses locker rooms, a training room, offices, a weapons storage vault and a classified vault. Entry into the building is through one of two doorways located on the south side of the building, adjacent to the parking lot, and through an additional doorway on the west side which primarily provides access between the physical training room and an outside quarter mile track. The building is within a Property Protection Area delineated by a standard 8 foot high chain link fence line. A fence line gate is open during operational hours. After hours, entry is made through a two door entry control portal under the control of a security inspector located at the Sandia Headquarters Control Center.

DOE Requirements

The shell of the building defines the boundary of a Limited Area. DOE regulations require Limited Area access control systems to:

1. verify the identity of persons authorized access at the area entrance to the area;
2. maintain a visitor log and;
3. provide for a regularly applied test and maintenance program for security related subsystems and components

This work was supported by the United States Department of Energy under Contract DE-AC04-94AL85000.

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Other requirements address searches for contraband of individuals and vehicles, both of which will not be handled by this system. Searches of authorized persons and visitors are currently performed by a random selection system.

Operational Requirements

The system is required to provide an enhanced level of security without greatly increasing the level of user or administrative interaction. It was understood that some administrative overhead would be necessary. It should, however, be kept to a minimum. Operational requirements are relatively simple and straight forward. Primary requirements are as follows:

1. Enrollment into the system will be performed by a receptionist or administrative assistant.
2. Enrollment must be relatively quick (less than 5 minutes).
3. Enrollment should support self selection of personal identification number (PIN).
4. Use of the system should not impose unreasonable requirements or delays to the staff.
5. A system over-ride should exist in case of technical problems.
6. Entry will be controlled; exit will be free.

2.0 SYSTEM DESCRIPTION

Pre-Upgrade

Access into building 956 was originally controlled by an electronic lock activated by entering a single numeric code at keypads at each door. The single access code was shared by everyone with authorization to enter the building. This system is analogous to everyone having a copy of the same key. Though the system provided positive control, it did not provide unique identification. This feature was needed in order to change the building from a Property Protection Area to a Limited Area.

This system worked well in that it had few, if any, technical and operational problems. The users accepted the need to memorize new access codes every time policy dictated the code be changed. There were few mechanical problems and maintenance was simple and straight forward.

Post-Upgrade

During the summer of 1993, building 956 was upgraded to allow the entire building, with the exception of the exercise room, to become a Limited Area. The upgrade replaced the keypads with Recognition Systems Inc. HandKey, hand geometry identity verifiers. Three HandKeys were used. One each at the two main south doors and another at an interior door between the exercise room and the rest of the building. The HandKeys were standard devices except that the manufacturer provided a custom program which provided score data. Score data was needed to better quantify performance. The two HandKeys used at the main doors were mounted on the exterior of the building. Environmental housings were used with these two verifiers.

3.0 TEST OBJECTIVES

The objective in this field study was to determine differences in performance between the laboratory and field installations as well as to develop tools which could be used by operating sites and at future field tests to gauge the effectiveness of biometric access control equipment. During the field study, efforts were made to identify metrics and methods of testing which could characterize biometric systems. The metrics which were developed were intended to be generic and thus applicable to a wide range of biometric verifiers.

4.0 OPERATIONAL DESCRIPTION

Enrollment

Those persons who had previously been on the access list to building 956, were invited to enroll into the system prior to the system being placed on line. The host computer, running RSI's HandNet software, and a single HandKey was used for this pre-enrollment. Users were trained in the use of the HandKeys and were allowed to select their own PIN. The pre-enrollment database was broadcast to all readers once the system became operational.

New persons requesting access would first be verified by the system administrator as having official business. The new enrollee would then be entered into the database and enrolled during their first visit. During this visit, the enrollee would be trained on the use of the system.

Several generic user groups for the system were defined. Users were classified according to access schedules and authority levels. The various groups were as follows:

Managers/Administrators - 24 hr. access/full configuration authority
Maintenance - 40 hr. week access/full configuration authority
Security - 24 hr. access/no configuration authority
Staff - 40 hr. week access/no configuration authority

These levels worked well for the testing, however, other security level definitions may be warranted in a larger operational system.

Dis-Enrollment

When a user's need to access the building expires, the system administrator would remove the user from the database. This was done with the 'Remove User' utility in the HandNet software. The command is broadcast to all readers on the net and the user is henceforth denied access.

5.0 TEST RESULTS

Baseline Data

To establish a baseline, the system was allowed to run undisturbed for a period of four weeks after an initial setup and adjustment period. During this time over 6000 transactions were recorded. These transactions made up the baseline database. The transaction data collected included name, time, PIN, entry point, and score.

Variability of Reader

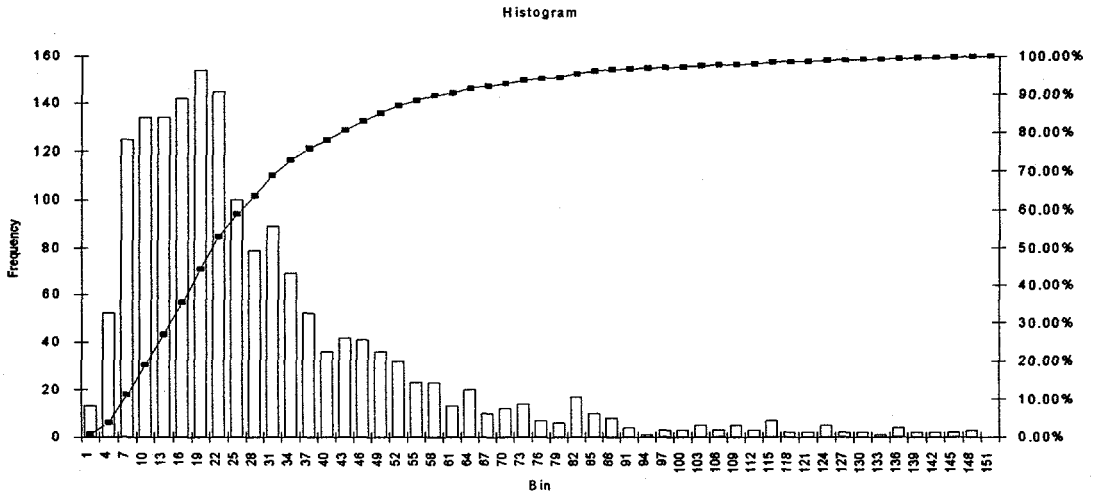
The distribution of individual and system wide scores is an indicator which speaks to the repeatability of the biometric readings. Low numbers indicate an invariant system. This value however includes variability due to both the device and the method used by the individual in presenting the biometric. To determine these variabilities, we needed to isolate the various sources of error. We first looked at the variability of single readers. To measure this, a computer was programmed to collect scores from individual readers presented with a fixed reference mechanical hand. At least 200 measurements were taken from each reader. Since the 'hand' is a mechanical surrogate and all other conditions were controlled, the readings should have been identical. The difference in readings is a measure of the exactness of the device. This test was performed at various temperatures. The variance of the devices in the test system were consistently no greater than 1 score point of the expected value. There was no noticeable effect due to temperature.

Variability Between Readers

Another potential error factor that is meaningful in evaluation of biometric devices is the variability between devices. The three HandKey readers in the system were each presented the same fixed mechanical hand and the recorded templates were compared. The difference in scores varied a maximum of 5 points. This indicates that the HandKey readers are fairly uniform, one to another, in making a biometric reading.

Variability of Presented Biometric

One final error factor which may affect an individual's score at the HandKey reader is a person's ability to consistently present their hand the same way. We found that most of the score variance is due to the differences in the way a person presents their hand. To determine this error, a 'C' program was written to collect and calculate score deviations. After loading the transaction database, this program sorts the transactions by PIN and then uses the scores to calculate individual score variations. The individual deviation values are then used to calculate a system deviation. The average score was found to be 34 with an average standard deviation of 20. The following chart of this data displays the wide possible scores.



Baseline distribution of scores (prior to any system optimization, system installed as per manufacturer's recommendations); average = 34, mean = 25, standard deviation = 20.

False Reject Rate

One final baseline metric recorded was the system's false reject rate. This error rate is defined as the total number of rejections of authorized persons divided by the total number of transactions. As can be imagined, this error rate will vary depending on

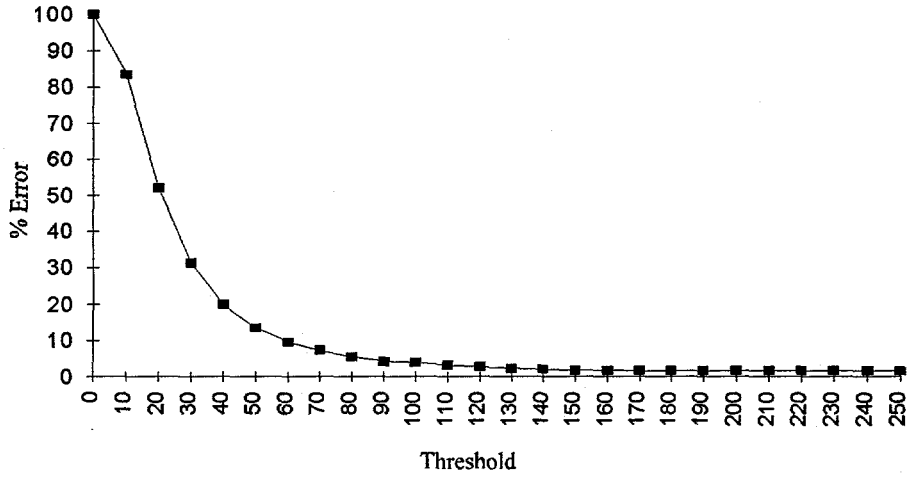
the system threshold, or sensitivity. The lower the threshold (increased sensitivity), the higher the error.

A 'C' program was also written to determine this error rate. The program shifts through the database and using a set of rules, identifies first, second, and third attempt transactions. These transactions along with the score were used to compute the error rates at the various thresholds. The program produces a file containing the matrix of threshold versus error rate for first attempt transactions. This file was then used to plot the results using the Microsoft Excel program.

The results of the baseline system do not compare very well with the data generated during previous laboratory testing. The results are tabulated below.

False Rejection Failure Rates from transaction of data of all three doors:

First Attempt:	02.23%	(134 of 6000)
Second Attempt:	43.06%	(31 of 72)
Third Attempt:	57.89%	(11 of 19)



False Reject Error rate of base line configuration developed from data of the two exterior doors.

This data seems to indicate that when a person is first rejected the reason for the rejection is most likely due to some factor other than sloppy presentation of the requester's hand. We have seen in our laboratory testing that after being rejected a person is much more careful on their second and third attempts and thus the error rates go down for these transactions. A possible cause for these higher numbers is forgetting or incorrectly entering the PIN, or some environmental factor not present in the laboratory testing. Direct sunlight and blowing dust and sand being the most probable.

The data shows that out of the 134 rejected first attempts only 72 persons made a second attempt and from that group 31 were rejected a second time. From this group of 31, only 19 attempted it again. Why they didn't all try again was a point of some concern and was investigated further. It was found that frequently people came to the building in groups and if one person failed in their first or second attempt another in the group would try. This was reasonable given the security level of the building. However another practice uncovered was not reasonable and possibly detrimental to the security system. On some occasions, when a security police officer was not granted access, the officers would use their key to unlock the door. This practice is not desired in an operational security system since the alarm at the door is not masked and so is enunciated at the security control center requiring a response.

System Optimization

Traditionally, a single system wide threshold is set for all users. There is the capability to change thresholds on an individual basis but this is cumbersome and is usually only done when an individual is experiencing difficulties in gaining access. In these limited cases the threshold can be increased. For this field application, optimization was based on the fact that after an initial period, the transaction database holds the required information to individually set thresholds based on some statistical criteria. The system could have been optimized in two ways. First, it could be optimized for high security by setting individual thresholds to their lowest possible value based on some criteria. Secondly, the system could be optimized for user acceptance, meaning high thresholds. Either of these methods could be adjusted to suit site specific needs.

High Security Type of Optimization

This optimization strives to determine the lowest individual threshold which would ensure a near zero false accept error rate. The optimization drives down each individual's threshold based on statistical criteria. The process begins by collecting individual user scores and then calculating the mean and the standard deviation. The standard deviation, or sigma point, can then be used to calculate the threshold of the user. Thresholds are calculated by the following formula:

threshold = mean + 3 sigma

At this threshold setting the user is assured a high probability of consistent acceptance. The three sigma point is a variable which could be adjusted by the security manager to meet specific needs. Additional criteria could set a maximum threshold, above which no individual threshold would be set.

The tool which performs this calculation is a C program named 'OPTI.C'. This program collects the individual scores, sorted by PIN, from the standard 'REPORT' file. It then calculates the three sigma point and writes the value to a separate file. This file is in turn used by the optimization program to update the thresholds.

User Ease Optimization

The system may also be optimized by driving individual thresholds upwards. This can be done by comparing individual templates, one against the other. By doing this, each template's closest match is identified. These template pairs are called cohorts. Once the cohorts are found, the scores which would result from the comparison of the two templates is known. This score represents the score which would be expected for a user compared to that user's closest match. In the baseline study we found that the majority of persons in the database had cohorts greater than 150 score points away and many were over 200 score points away. If these scores were used as individual thresholds, then one could expect that no one person in the database could get in for any other. In addition, the ease of use would be increased because for the majority of persons their thresholds would be increased.

This method does not consider persons not enrolled in the system. The system thus relies on the low probability of someone randomly selecting a valid PIN of a person to whom they would match. Admittedly this method for optimization is not a high security option, but it does offer advantages for applications which only require positive identification of individual employees as one individual could not pass for another. Optimization performed in this manner was not tested in the field environment.

Compromised Optimization

By combining elements of the previous two methods, a good compromised optimization approach could be developed. The elements which may be combined are the individual sigma points and the knowledge of the cohorts. The sigma points define the highest threshold while the cohort score defines the lowest. Where the sigma point is less than the cohort score the sigma point is used as the threshold. In those instances where the cohort score is lower than the sigma score, some other value, based on the individual standard deviation, is used.

High Security Test Series

The test series looked at the high security mode of operation. All baseline data was saved to an archival disk. The false accept rate of the data was then calculated and found to be 0.59%. A new user database was created with calculated individual thresholds. This database was down-loaded to the three readers and the system was allowed to run for 2 months. The following table compares the baseline performance with the performance using the high security optimization.

	Baseline	**Optimized**
False Accept	0.59%	0.42%
False Reject	2.23%	5.20%

The optimized system improves the probability of detection (Pd) by 35%. This increased performance is obtained at the cost of three additional people out of 100 being falsely rejected during their first attempt at gaining access. This value may even be high, as the data from which it was computed contained a series of transactions which occurred after a dust storm which greatly increased the false reject error. It is probable that a better maintained system would have a lower false reject rate.

Statistical Process Control (SPC)

By collecting all transaction data, one can perform standard statistical control processes and determine how the individual devices and/or system is performing as well as how the performance is affected by varying setup parameters. SPC is very powerful and does not impose additional data collection requirements. The data that is used is that which would normally be collected, so the benefits of SPC can be achieved for little extra effort. For a more detailed discussion, the reader is referred to laboratory report "*Standardized Testing of Hand Geometry Identification Verifiers*", by Dale Murray.

6.0 TOOLS

The following tools were developed to assist in the analysis and optimization of this HandKey system.

The tools are Microsoft Excel macros and DOS executables written in the C programming language. Software listings are available on request.

Standard Report File

The RSI control software, HandNet, provides a report generation option under the Report menu. A standard report format is used to collect all available data. This format is stored as a definition file called 'STDRPT'. When STDRPT is used to generate a report file all information needed for the Excel macros and the C executables, will be included.

The standard report will be written to the A: diskette drive, under the file name 'REPORT'. A diskette was used so that further analysis could be completed at a different location allowing the access control computer to return to logging and controlling transactions and activities.

Microsoft Excel Macros

The Excel macros were written for Excel version 4.0. There are two macros. The first, titled INTER.XLM, reads in a standard report file, STDRPT, and formats it for further analysis. The formatted file is saved as BASE.DAT. The second Excel macro opens BASE.DAT and sorts the transactions by PIN. It then calculates the average score and standard deviation for each PIN. This file is called PINS.XLM. The third macro is called ERRORATE.XLM. This macro also uses the file BASE.DAT, but it calculates the false reject rate of the system.

It was learned that Microsoft was discontinuing the Excel version 4.0 macro language in favor of Visual Basic. For this reason the macro utilities were converted to 'C' language programs. The macros are mentioned here for the sake of completeness.

C Programs

HANDANLZ uses the standard report generated by Recognition System HandNet software and identifies transactions as first, second, or third attempts. It then calculates the rejection rate for each category of transaction. (This is of greatest use for determination of the false rejection rate, when every transaction should be accepted.) This data is displayed on the screen and written to a summary file, usually with extension .SUM. The program also creates an intermediate file, with the extension .1ST, consisting of all first-attempt transactions with useable scores.

HNDTHRES takes as input the intermediate file generated by HANDANLZ and calculates the first-attempt rejection rate that would have occurred at thresholds between 0 and 250. It writes a table of this data to a data file, usually with the extension .TAB.

HAND is the front end for the two analysis programs. It prompts the user for the name of the initial data file and a step size for the threshold analysis. It then constructs output filenames by adding the proper extensions to the original filename, checks to make sure the files are viable, and calls HANDANLZ and HNDTHRES in sequence. This is the recommended way to perform the analysis.

7.0 CONCLUSION

In testing a biometric in actual field conditions, more was learned about the limitations of the system than possible under laboratory conditions. It was learned that direct solar light interferes with the device in making a reading, this increasing the false reject rate. Dust and sand blown up into the device also interferes with the system and could possibly be a vulnerability. User acceptance was not as good as was seen in the lab due to the rapid wear and accumulation of dirt and grime that exterior units experienced. Finally, it was determined that a proper maintenance program is very important in order to keep error rates down. Laboratory performance and field performance were found to vary greatly.

The optimization process appears to be a very good solution for improving the performance of a HandKey system. The optimization provides the site security manager another dimension within which to fine tune an entry control system. Greater ease of use, higher security, or a compromised system can be developed with the optimization techniques presented here. This coupled with the ability to constantly monitor the performance of the system and provide timely maintenance, promises greater overall performance which translates into ease of use and increased probabilities of detection of unauthorized entry attempts.

The ability to use existing transaction data to monitor a system's performance is very exciting. As the data is 'free', the benefits come at a low cost. When the savings from not having to periodically test the system are included, the benefits are extremely cost effective.