Appendix A

BIOMETRICS: A TECHNICAL PRIMER

This appendix expands on information presented in Chapter Two. It begins with a definition of biometrics and related terms, and then describes the steps in the biometric authentication process, reviews issues of template management and storage, and addresses testing issues. The appendix concludes with a brief review of mainstream biometric applications.[1]

DEFINITIONS

A biometric is any *measurable, robust, distinctive* physical characteristic or personal trait that can be used to *identify,* or *verify the claimed identity* of, an individual. Biometric authentication, in the context of this report, refers to automated methods of identifying, or verifying the identity of, a *living person*.

The italicized terms above require explanation. *Measurable* means that the characteristic or trait can be easily presented to a sensor and converted into a quantifiable, digital format. This allows for the automated matching process to occur in a matter of seconds.

The *robustness* of a biometric is a measure of the extent to which the characteristic or trait is subject to significant changes over time.

[1]This primer does not cover standards for interoperability or so-called "plug and play" applications because this subject is tangential to this RAND project. In researching and writing this appendix, the authors relied heavily on the following sources: Hawkes and Hefferman (1999); Newton and Rubenson (1999); and Wayman (1999c, 2000). See also Dunn (1998) and generally, Jain, Bolle, and Pankanti (1998).

These changes can occur as a result of age, injury, illness, occupational use, or chemical exposure. A highly robust biometric does not change significantly over time. A less robust biometric does. For example, the iris, which changes very little over a person's lifetime, is more robust than a voice.

Distinctiveness is a measure of the variations or differences in the biometric pattern among the general population. The higher the degree of distinctiveness, the more unique the identifier. The highest degree of distinctiveness implies a unique identifier. A low degree of distinctiveness indicates a biometric pattern found frequently in the general population. The iris and the retina have higher degrees of distinctiveness than hand or finger geometry. The application helps determine the degree of robustness and distinctiveness required.

Living person distinguishes biometric authentication from forensics, which does not involve real-time identification of a living individual.

IDENTIFICATION VERSUS VERIFICATION

Identification and verification differ significantly. With identification, the biometric system asks and attempts to answer the question, "Who is X?" In an *identification application*, the biometric device reads a sample and compares that sample against every template in the database. This is called a "one-to-many" search (1:N). The device will either make a match and subsequently identify the person or it will not make a match and not be able to identify the person.

Verification is when the biometric system asks and attempts to answer the question, "Is this X?" after the user claims to be X. In a *verification application*, the biometric device requires input from the user, at which time the user claims his identity via a password, token, or user name (or any combination of the three). This user input points the device to a template in the database. The device also requires a biometric sample from the user. It then compares the sample to or against the user-defined template. This is called a "one-to-one" search (1:1). The device will either find or fail to find a match between the two.

Identification applications require a highly robust and distinctive biometric, otherwise the error rates falsely matching and falsely

nonmatching users' samples against templates cause security problems and inhibit convenience. Identification applications are common where the end-user wants to identify criminals (immigration, law enforcement, etc.) or other "wolves in sheep's clothing." Other types of applications may use a verification process.[2] In many ways, deciding whether to use identification or verification requires a trade-off: the end-user's needs for security versus convenience.

In sum, biometric authentication is used in two ways: to prove who you are or who you claim you are and to prove who you are not (e.g., to resolve a case of mistaken identity).

APPROACHES TO AUTHENTICATION

In general, there are three approaches to authentication. In order of most secure and convenient to least secure and convenient, they are as follows:

- Something you are—a biometric.
- Something you know—PIN, password.
- Something you have—key, token, card.

Any combination of these approaches further heightens security. Requiring all three for an application provides the highest form of security.[3]

THREE BASIC ELEMENTS TO ALL BIOMETRIC SYSTEMS

All biometric systems consist of three basic elements:

- Enrollment, or the process of collecting biometric samples from an individual, known as the enrollee, and the subsequent generation of his template.
- Templates, or the data representing the enrollee's biometric.

---

[2]See, e.g., Appendix B, Program Reports, Fort Sill Biometrically Protected Smart Card.

[3]Security also depends on other factors, such as the care taken to apply security measures properly, insofar as safeguarding tokens and passwords and ensuring that transmissions of biometric data are adequately protected.

- Matching, or the process of comparing a live biometric sample against one or many templates in the system's database.

Enrollment

Enrollment is the crucial first stage for biometric authentication because enrollment generates a template that will be used for all subsequent matching. Typically, the device takes three samples of the same biometric and averages them to produce an enrollment template. Enrollment is complicated by the dependence of the performance of many biometric systems on the users' familiarity with the biometric device because enrollment is usually the first time the user is exposed to the device.

Environmental conditions also affect enrollment. Enrollment should take place under conditions similar to those expected during the routine matching process. For example, if voice verification is used in an environment where there is background noise, the system's ability to match voices to enrolled templates depends on capturing these templates in the same environment.[4]

In addition to user and environmental issues, biometrics themselves change over time. Many biometric systems account for these changes by continuously averaging. Templates are averaged and updated each time the user attempts authentication.

Templates

As the data representing the enrollee's biometric, templates are created by the biometric device. The device uses a proprietary algorithm to extract "features" appropriate to that biometric from the enrollee's samples. Templates are only a record of distinguishing features, sometimes called minutiae points, of a person's biometric characteristic or trait. For example, templates are not an image or record of the actual fingerprint or voice.[5] In basic terms, templates

[4]The system's ability to match the sample to the enrolled template is sometimes referred to as the biometric's reliability.

[5]Image files of fingerprints may be of interest to the Army because of their law enforcement applications. In the case of fingerprints, the Army may want to keep

are numerical representations of key points taken from a person's body.

The template is usually small in terms of computer memory use, and this allows for quick processing, which is a hallmark of biometric authentication. The template must be stored somewhere so that subsequent templates, created when a user tries to access the system using a sensor, can be compared. Some biometric experts claim it is impossible to reverse-engineer, or recreate, a person's print or image from the biometric template.

Matching

Matching is the comparison of two templates, the template produced at the time of enrollment (or at previous sessions, if there is continuous updating) with the one produced "on the spot" as a user tries to gain access by providing a biometric via a sensor.

There are three ways a match can fail:

- Failure to enroll.
- False match.
- False nonmatch.

Failure to enroll (or acquire) is the failure of the technology to extract distinguishing features appropriate to that technology. For example, a small percentage of the population fails to enroll in fingerprint-based biometric authentication systems. Two reasons account for this failure: the individual's fingerprints are not distinctive enough to be picked up by the system, or the distinguishing characteristics of the individual's fingerprints have been altered because of the individual's age or occupation, e.g., an elderly bricklayer.

---

both electronic image files of the fingerprint as well as the biometric templates. The image files are too large to be used for biometric applications but would be useful for forensic purposes. Moreover, the Army might want to store image files to give it greater technical flexibility. For example, if the Army did not keep image files of enrollees, it might have to physically reenroll each individual if the Army decided to change to a different proprietary biometric system. Image files are also known as raw data or the *corpus*.

In addition, the possibility of a false match (FM) or a false nonmatch (FNM) exists. These two terms are frequently misnomered "false acceptance" and "false rejection," respectively, but these terms are application-dependent in meaning. FM and FNM are application-neutral terms to describe the matching process between a live sample and a biometric template.

A false match occurs when a sample is incorrectly matched to a template in the database (i.e., an imposter is accepted). A false nonmatch occurs when a sample is incorrectly not matched to a truly matching template in the database (i.e., a legitimate match is denied). Rates for FM and FNM are calculated and used to make tradeoffs between security and convenience. For example, a heavy security emphasis errs on the side of denying legitimate matches and does not tolerate acceptance of imposters. A heavy emphasis on user convenience results in little tolerance for denying legitimate matches but will tolerate some acceptance of imposters.

TEMPLATE MANAGEMENT—STORAGE AND SECURITY

Template management is critically linked to privacy, security, and convenience issues. All biometric authentication systems face a common issue: Biometric templates must be stored somewhere. Templates must be protected to prevent identity fraud and to protect the privacy of users. A major concern is what additional information will be stored about each user along with his biometric template.

Possible locations for template storage include

- the biometric device itself,
- a central computer that is remotely accessed,
- a plastic card or token via a bar code or magnetic stripe,
- Radio Frequency Identification Device cards and tags,
- optical memory cards,
- Personal Computer Memory Card International Association cards, and
- smart cards.

In general, transmitting biometric data over communications lines reduces system security because the data become vulnerable to the same interception or tampering possible when any data is sent "over the wire." Biometrics are more secure when stored under the control of the authorized user, such as on a smart card, and used in verification applications.

Smart cards are the size of credit cards and have a microchip or microprocessor chip embedded in them. The chip stores electronic data that can be protected using biometrics. There are two types of smart cards: contact and contactless smart cards. A contact smart card must be inserted into a smart card reader to be used. A contactless smart card only has to be placed near an antenna to carry out a transaction.[6]

Another security issue for template database storage is whether the database will have a unique use or if it will be used for multiple security uses. For example, a facilities manager might use a fingerprint reader for physical access control to the building. The manager might also want to use the same fingerprint template database for his employees to access their computer network. Should the manager use separate databases for these different uses, or is he willing to risk accessing employee fingerprints from a remote location for multiple purposes?

In general, verification applications provide more security than identification applications because a biometric and at least one other piece of input (e.g., PIN, password, token, user name) are required to match a template. Verification provides a user with more control over his data and over the process when the template is stored only on a card. That is, such a system would not allow for clandestine, or involuntary, capture of biometric data because the individual would know if he were providing the card. Because the search only seeks a match against one template in the database, verification applications require less processing time and memory. Thus, they are less expensive than identification applications.

Additional security features can be incorporated into biometric systems to detect a "wolf," or unauthorized user. For example, a

[6]For a detailed discussion of smart cards, see Ratha and Bolle (1999).

"liveliness test" is a method of measuring if the biometric sample is being read from a live person versus a faux body part or body part of a dead person. Liveliness tests are done in many ways. The device can look for such things as heat, heartbeat, or electrical capacitance.[7] Other security features include encryption of biometric data and the use of sequence numbers in template transmission. A template with such a number out of sequence suggests unauthorized use.

MAINSTREAM BIOMETRICS AND THEIR APPLICATIONS

While there are many possible biometrics, at least eight mainstream biometric authentication technologies have been deployed or pilot-tested in applications in the public and private sectors.[8] These are

- fingerprint,
- hand/finger geometry,
- facial recognition,
- voice recognition,
- iris scan,
- retinal scan,
- dynamic signature verification, and
- keystroke dynamics.

Fingerprint

The fingerprint biometric is an automated digital version of the old ink-and-paper method used for more than a century for identification, primarily by law enforcement agencies. The biometric device involves users placing their finger on a platen for the print to be read. The minutiae are then extracted by the vendor's algorithm, which

---

[7]Electrical capacitance has proved to be the best and least reproducible method for effectively identifying a live person.

[8]For a detailed discussion of these mainstream biometrics, see Jain, Bolle, and Pankanti (1999).

also makes a fingerprint pattern analysis. Fingerprint template sizes are typically 50 to 1,000 bytes.

Fingerprint biometrics currently have three main application arenas: large-scale Automated Finger Imaging Systems (AFIS) generally used for law enforcement purposes, fraud prevention in entitlement programs, and physical and computer access.

Hand/Finger Geometry

Hand or finger geometry is an automated measurement of many dimensions of the hand and fingers. Neither of these methods takes actual prints of the palm or fingers. Only the spatial geometry is examined as the user puts his hand on the sensor's surface and uses guiding poles between the fingers to properly place the hand and initiate the reading. Hand geometry templates are typically 9 bytes, and finger geometry templates are 20 to 25 bytes. Finger geometry usually measures two or three fingers. During the 1996 Summer Olympics, hand geometry secured the athlete's dormitories at Georgia Tech. Hand geometry is a well-developed technology that has been thoroughly field-tested and is easily accepted by users.

Facial Recognition

Facial recognition records the spatial geometry of distinguishing features of the face. Different vendors use different methods of facial recognition, however, all focus on measures of key features. Facial recognition templates are typically 83 to 1,000 bytes. Facial recognition technologies can encounter performance problems stemming from such factors as noncooperative behavior of the user, lighting, and other environmental variables. Facial recognition has been used in projects to identify card counters in casinos, shoplifters in stores, criminals in targeted urban areas, and terrorists overseas.

Voice Recognition

Voice or speaker recognition uses vocal characteristics to identify individuals using a pass-phrase. Voice recognition can be affected by such environmental factors as background noise. Additionally it is unclear whether the technologies actually recognize the voice or just

the pronunciation of the pass-phrase (password) used. This technology has been the focus of considerable efforts on the part of the telecommunications industry and NSA, which continue to work on improving reliability. A telephone or microphone can serve as a sensor, which makes it a relatively cheap and easily deployable technology.

Iris Scan

Iris scanning measures the iris pattern in the colored part of the eye, although the iris color has nothing to do with the biometric. Iris patterns are formed randomly. As a result, the iris patterns in your left and right eyes are different, and so are the iris patterns of identical twins. Iris scan templates are typically around 256 bytes. Iris scanning can be used quickly for both identification and verification applications because of its large number of degrees of freedom. Current pilot programs and applications include ATMs ("Eye-TMs"), grocery stores (for checking out), and the Charlotte/Douglas International Airport (physical access). During the Winter Olympics in Nagano, Japan, an iris scanning identification system controlled access to the rifles used in the biathlon.

Retinal Scan

Retinal scans measure the blood vessel patterns in the back of the eye. Retinal scan templates are typically 40 to 96 bytes. Because users perceive the technology to be somewhat intrusive, retinal scanning has not gained popularity with end-users. The device involves a light source shined into the eye of a user who must be standing very still within inches of the device. Because the retina can change with certain medical conditions, such as pregnancy, high blood pressure, and AIDS, this biometric might have the potential to reveal more information than just an individual's identity.

Dynamic Signature Verification

Dynamic signature verification is an automated method of examining an individual's signature. This technology examines such dynamics as speed, direction, and pressure of writing; the time that

the stylus is in and out of contact with the "paper"; the total time taken to make the signature; and where the stylus is raised from and lowered onto the "paper." Dynamic signature verification templates are typically 50 to 300 bytes.

Keystroke Dynamics

Keystroke dynamics is an automated method of examining an individual's keystrokes on a keyboard. This technology examines such dynamics as speed and pressure, the total time of typing a particular password, and the time a user takes between hitting certain keys. This technology's algorithms are still being developed to improve robustness and distinctiveness. One potentially useful application that may emerge is computer access, where this biometric could be used to verify the computer user's identity continuously.

BIOMETRIC APPLICATIONS

Most biometric applications fall into one of nine general categories:

- Financial services (e.g., ATMs and kiosks).
- Immigration and border control (e.g., points of entry, precleared frequent travelers, passport and visa issuance, asylum cases).
- Social services (e.g., fraud prevention in entitlement programs).
- Health care (e.g., security measure for privacy of medical records).
- Physical access control (e.g., institutional, government, and residential).
- Time and attendance (e.g., replacement of time punchcard).
- Computer security (e.g., personal computer access, network access, Internet use, e-commerce, e-mail, encryption).
- Telecommunications (e.g., mobile phones, call center technology, phone cards, televised shopping).
- Law enforcement (e.g., criminal investigation, national ID, driver's license, correctional institutions/prisons, home confinement, smart gun).

Classifying Biometric Applications

Biometric applications may be classified in many different ways. James Wayman of the National Biometric Test Center suggests the following seven categories for classifying biometric applications, explained below.

1. overt or clandestine
2. cooperative or noncooperative
3. habituated or nonhabituated
4. supervised or nonsupervised
5. standard or nonstandard environment
6. closed or open system
7. public or private.

Overt versus clandestine capture of a biometric sample refers to the user's awareness that he is participating in biometric authentication.[9] Facial recognition is an example of a biometric that can be used for clandestine identification of individuals. Most uses of biometrics are overt, because users' active participation ensures performance and lower error rates. Verification applications are nearly always overt.

Cooperative versus noncooperative applications refer to the behavior that is in the best interest of the "wolf." Is it in the interest of "wolves" to match or to not match a template in the database? Which is to the "wolf's" benefit? This is important in planning a security system with biometrics. No perfect biometric system exists; every system can be tricked into falsely not matching one's sample and template—some more easily than others. It is also possible to trick a biometric device into falsely matching your sample against a template, but it could be argued that this requires more work and a sophisticated hacker to make a model of the biometric sample. One way to strengthen security in a cooperative application is to require a password or token along with a biometric, so that the "wolf" must

[9] James Wayman used "covert" instead of "clandestine."

match one specific template and is not allowed to exploit the entire database for his gain.

To gain access to a computer, a "wolf" would want to be cooperative. To attempt to foil an INS database consisting of illegal border crossing recidivists, a "wolf" (recidivist) would be noncooperative.

Habituated versus nonhabituated use of a biometric system refers to how often the users interface with the biometric device. This is significant because the user's familiarity with the device affects its performance. Depending on which type of application is chosen, the end-user may need to utilize a biometric that is highly robust. As examples, use of fingerprints for computer or network access is a habituated use; use of fingerprints on a driver's license, which is updated after several years, is a nonhabituated use. Even "habituated" applications are "nonhabituated" during their first week or so of operation or until the users adjust to using the system.

Supervised versus nonsupervised applications refer to whether supervision (e.g., a security officer) is a resource available to the end-user's security system. Do users need to be instructed on how to use the device (many new users or nonhabituated users) or to be supervised to ensure they are being properly sampled (such as border crossing situations with the problem of recidivists or other noncooperative applications)? Or is the application made for increased convenience, such as at an ATM? The process of enrollment nearly always requires supervision.

Standard versus nonstandard environments are generally a dichotomy between indoors versus outdoors. A standard environment is optimal for a biometric system and matching performance. A nonstandard environment may present variables that would create false nonmatches. For example, a facial recognition template depends, in part, on the lighting conditions when the "picture" (image) was taken. The variable lighting outdoors can cause false nonmatches. Some indoor situations may also be considered nonstandard environments.

Closed versus open systems refers to the number of uses of the template database now or potential uses in the future. Will the database have a unique use (closed), or will it be used for multiple security measures (open)? For example, a facilities manager might have his

employees use a fingerprint reader to enter a building. He might also want to use the same fingerprint template database for employees to log on to their computer network. Should they use separate databases for these different uses, or do they want to risk remotely accessing employee fingerprints for multiple purposes? Other examples are state driver's licenses and entitlement programs. A state may want to communicate with other states or other programs within the same state to eliminate fraud. This would be an open system, in which standard formats of data and compression would be required to exchange and compare information.

Public or private applications refer to the users and their relationship to system management. Examples of users of public applications include customers and entitlement recipients. Users of private applications include employees of business or government. The users' attitudes toward biometric devices and management's approach will vary depending on whether the application is public or private. Once again, users' attitudes toward the device will affect the performance of the biometric system.

It should be noted here that performance figures and error rates from vendor testing are unreliable for many reasons. Part of the problem is that to test the distinctiveness of a biometric, anywhere from thousands to millions of people are needed to test theories of how "unique" a particular identifier is. To acquire samples over any amount of time in any number of contexts from this number of people would be impossible, and to do this same testing for the many variables in each type of application is in most cases impossible and in the others too costly if it were possible. Operational and pilot testing is the only reasonable method to test a system. Additionally, vendor and scientific laboratory testing generally present only one scenario of biometric application: overt, cooperative, habituated, supervised, standard, closed, and private (Newton and Webb, 1999).

SALIENT CHARACTERISTICS OF MAINSTREAM BIOMETRICS

Table A.1 compares the eight mainstream biometrics in terms of a number of characteristics, ranging from how robust and distinctive

Table A.1

Comparison of Mainstream Biometrics[10]

Biometric	Identify versus Verify	How Robust	How Distinctive	How Intrusive
Fingerprint	Either	Moderate	High	Touching
Hand/Finger Geometry	Verify	Moderate	Low	Touching
Facial Recognition	Either	Moderate	Moderate	12+ inches
Voice Recognition	Verify	Moderate	Low	Remote
Iris Scan	Either	High	High	12+ inches
Retinal Scan	Either	High	High	1–2 inches
Dynamic Signature Verification	Verify	Low	Moderate	Touching
Keystroke Dynamics	Verify	Low	Low	Touching

they are to what they can be used for (i.e., identification or verification or verification alone) (Newton and Webb, 1999). This table is an attempt to assist the reader in categorizing biometrics along important dimensions. Because this industry is still working to establish comprehensive standards and the technology is changing rapidly, however, it is difficult to make assessments with which everyone would agree. The table represents an assessment based on discussions with technologists, vendors, and program managers. The table is not intended to be an aid to those in the market for biometrics, rather it is a guide for the uninitiated.

When comparing ways of using biometrics, half can be used for either identification or verification, and the rest can only be used for verification. In particular, hand geometry has only been used for verification applications, such as physical access control and time and attendance verification. In addition, voice recognition, because of the need for enrollment and matching using a pass-phrase, is typically used for verification only.

There is considerable variability in terms of robustness and distinctiveness. Fingerprinting is moderately robust, and, although it is

[10]The authors compiled Table A.1 from various sources at the SJB Biometrics 99 Workshop, November 9–11, 1999, including Hawkes and Hefferman (1999). See also Jain, Bolle, and Pankanti (1998).

distinctive, a small percentage of the population has unusable prints, usually because of age, genetics, injury, occupation, exposure to chemicals, or other occupational hazards. Hand/finger geometry is moderate on the distinctiveness scale, but it is not very robust, while facial recognition is neither highly robust nor distinctive. As for voice recognition, assuming the voice and not the pronunciation is what is being measured, this biometric is moderately robust and distinctive. Iris scans are both highly robust because they are not highly susceptible to day-to-day changes or damages and distinctive because they are randomly formed. Retinal scans are fairly robust and very distinctive. Finally, neither dynamic signature verification nor keystroke dynamics are particularly robust or distinctive.

As the table shows, the biometrics vary in terms of how intrusive they are, ranging from those biometrics that require touching to others that can recognize an individual from a distance.

BIOMETRIC RDT&E CAPABILITIES

Biometrics are an emerging technology in an emerging industry that does not yet have comprehensive standards. As a result, test and evaluation will be an important component of an Army biometrics program.

The U.S. NBTC at San Jose State University in California, researches application-specific testing of systems and develops statistical methodologies for such operational and scenario testing. NBTC, directed by James L. Wayman, is primarily funded through the NSA. NBTC does not research or test specific biometric products. Its work must be linked to an application.

Commercial vendors also evaluate biometric devices. However, vendor testing is not independent, and results are not always replicable by others, such as the NBTC or the National Physical Laboratory (NPL) in England. In general, the performance of systems tested in a lab declines when the system is field tested.

There are six basic types of testing for biometric systems:

1. algorithm verification,
2. operational,

3. scenario,
4. usability,
5. security, and
6. template quality.[11]

NBTC Director Wayman and other experts include a seventh type of testing that is more cognitive: development of mathematical and statistical methods for test design and evaluation of biometric systems. Testing related to types 1, 2, and 3 are measurements of error rates.

Each basic type of testing is discussed below.

Test 1. Testers evaluate algorithms used by a single device using a database of "standard" samples. Standard samples are neither too easy nor too difficult for matching, and they probably do not reflect the anomalies found in populations. The results of this testing determine which algorithms are "good" and which are "poor." Although these test are useful and repeatable, the results do not show real-life performance under real field conditions with real enrollee populations.

Test 2. Operational testing is typically used for evaluating pilot programs. It helps determine how the system will perform as a whole based on a specific application environment on the target population.

Test 3. Scenario evaluation is used to test the performance of multiple biometric systems in a modeled real-world application of interest to evaluate and compare performance across biometric devices. All devices are tested in the same environment on the same population. The results are repeatable if the modeled scenario experiment can be controlled and should show real-life performance if done accurately. This method of evaluation allows for comparison of devices of different types. Scenario evaluation can help end-users decide which specific biometric device will work best for their needs.

---

[11]The discussion of the six types draws from information provided by Tony Mansfield, NPL, during a November 1999 interview and presentation in London, England (Newton and Webb, 1999).

Test 4. Usability of a biometric is critical to success, especially in commercial applications, because users must be willing to participate in the system and because usability enhances performance. Usability evaluations seek answers to such questions as whether the device is user-friendly, what difficulties users have with the system (e.g., intrusiveness, correct placement of biometric on the sensor), how the users' difficulties can be overcome, how users' difficulties with the system affect performance (e.g., false nonmatch rate), and whether the system is acceptable to end-users.

Test 5. Security evaluations seek answers to such questions as whether the system can detect imposters (a sufficiently low false match rate, liveliness tests, other measures), whether or not the device can distinguish between lookalikes (e.g., twins), whether some templates are easy to crack (i.e., does the device form templates for indistinguishable features that are easy to duplicate?), where the system is vulnerable, whether the system can be bypassed or hacked into and if so where and how.

Test 6. Evaluation of image/template quality could be useful for making standards for images'/templates' maximum allowable distortion, resolution, and signal-to-noise ratio. Standards for template quality will foster the ability for data-sharing between system managers. Through technical evaluation of biometric technologies, engineers and scientists search for the measurement of the following parameters: false match rate (FMR, the rate that a sample is incorrectly matched to a template in the database), false nonmatch rate (FNMR, the rate that a sample is incorrectly not matched to a template in the database), percentage of false nonmatches stemming from inconsistencies in the partitioning process[12] (known as the binning error rate), percentage of the total database to be scanned on average for each search (known as the penetration coefficient), transaction time (for finding resultant match or nonmatch), and failure to enroll/acquire percentage (percentage of the general population for whom the technology will fail to extract distinguishing features).

---

[12]Partitioning templates into smaller groups increases searching efficiencies and is used in systems holding a large number of templates. Partitioning can be based on information contained within the biometric template or other information gathered at the time of enrollment, such as the user's name or gender.

Even though biometric systems vary greatly across both biometric type and vendor, biometric devices have five subsystems: data collection, transmission, signal processing, storage of templates, and decision. Data collection occurs at the human-machine (sensor) interface and includes the creation of the biometric template. Transmission refers to the communication of the biometric template between the sensor and the next subsystem (either signal processing or storage depending on whether the user is trying to match a template[s] or enroll). This may include compression and subsequent expansion, which may add noise to the biometric pattern. Signal processing is when the device extracts distinguishing features from the biometric pattern presented at the sensor for matching and compares it to the template(s) stored during enrollment. Storage of a biometric template occurs at enrollment. The decision process takes the score from the signal processing subsystem and decides if a match or nonmatch is found, based on the thresholds the end-user has put into the system.

One test environment cannot predict error rates for all applications. Errors that would affect biometric authentication devices potentially come from four different sources: variations in the biometric pattern, the presentation of the biometric to the sensor, the sensor, and the transmission process (including compression and expansion noise). Each of these factors is strongly tied to a specific application. Hence, results from laboratory testing (vendor or otherwise) are dependent on the testing scenario and cannot usefully predict errors in real-world uses that are different.

Because vendor and scientific laboratory testing generally presents only the overt, cooperative, habituated, supervised, standard, closed, and private scenario, it is impossible to extrapolate performance in different sets of circumstances—such as in a nonhabituated or nonsupervised programs.

To test the distinctiveness of a biometric, anywhere from thousands to millions of people are needed to test theories of how “unique” a particular identifier is and to make statistically significant conclusions about uniqueness. Biometrics also “age” or change over time. To acquire samples over any amount of time (from weeks to months or even longer) in any number of contexts from this number of people would be close to impossible, and to do this same testing for

the many variables in each type of application would be even more difficult and probably financially prohibitive.

To summarize, as James Wayman has explained, three major difficulties occur in testing biometric devices and systems: "[1] the dependence of measured error rates on the application classification, [2] the need for a large test population [that] adequately models the target population, and [3] the necessity for a time delay between enrollment and testing." (Wayman, 1999e.)

While expensive, operational field or pilot testing and scenario evaluations are the only reasonable methods to test a system for deployment fully and reliably. Laboratory testing could be used to evaluate algorithms on an initial pass/fail basis for a biometric device to pass minimum standards to be further tested operationally. An R&D lab may also undertake further development of mathematical and statistical methods for test design and evaluation of biometric systems. An RDT&E center could be a source of advice on biometric systems for agencies internal and external to the Army, including being the developers of the educational roll-out piece of a biometric program.

An RDT&E center will face the testing difficulties highlighted above. At the same time, it could be useful in targeting research, developing mathematical and statistical methods for test design and evaluation, and screening the algorithms initially.