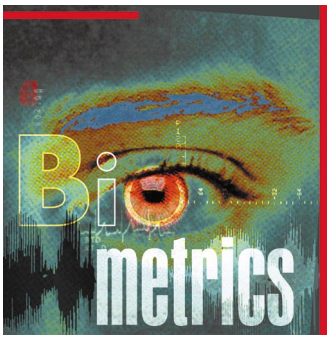
An Introduction to Evaluating Biometric Systems



How and where biometric systems are deployed will depend on their performance. Knowing what to ask and how to decipher the answers can help you evaluate the performance of these emerging technologies.

P. Jonathon Phillips
Alvin Martin
C.L. Wilson
Mark Przybocki
National Institute of Standards and Technology

On the basis of media hype alone, you might conclude that biometric passwords will soon replace their alphanumeric counterparts with versions that cannot be stolen, forgotten, lost, or given to another person. But what if the performance estimates of these systems are far more impressive than their actual performance?

To measure the real-life performance of biometric systems—and to understand their strengths and weaknesses better—we must understand the elements that comprise an ideal biometric system. In an ideal system

- all members of the population possess the characteristic that the biometric identifies, like irises or fingerprints;
- each biometric signature differs from all others in the controlled population;
- the biometric signatures don't vary under the conditions in which they are collected; and
- the system resists countermeasures.

Biometric-system evaluation quantifies how well biometric systems accommodate these properties. Typically, biometric evaluations require that an independent party design the evaluation, collect the test data, administer the test, and analyze the results.

We designed this article to provide you with sufficient information to know what questions to ask when evaluating a biometric system, and to assist you in determining if performance levels meet the requirements of your application. For example, if you plan to use a biometric to reduce—as opposed to eliminate—fraud, then a low-performance biometric system may be sufficient. On the other hand, completely replacing an existing security system with a biometric-based one may require a high-performance biometric system, or the required performance may be beyond what current technology can provide.

Here we focus on biometric applications that give the user some control over data acquisition. These applications recognize subjects from mug shots, passport photos, and scanned fingerprints. Examples not covered include recognition from surveillance photos or from latent fingerprints left at a crime scene.

Of the biometrics that meet these constraints, voice, face, and fingerprint systems have undergone the most study and testing—and therefore occupy the bulk of our discussion. While iris recognition has received much attention in the media lately, few independent evaluations of its effectiveness have been published.

PERFORMANCE STATISTICS

There are two kinds of biometric systems: *identification* and *verification*.

In identification systems, a biometric signature of an unknown person is presented to a system. The system compares the new biometric signature with a database of biometric signatures of known individuals. On the basis of the comparison, the system reports (or estimates) the identity of the unknown person from this database. Systems that rely on identification include those that the police use to identify people from fingerprints and mug shots. Civilian applications include those that check for multiple applications by the same person for welfare benefits and driver's licenses.

In verification systems, a user presents a biometric signature and a claim that a particular identity belongs to the biometric signature. The algorithm either accepts

or rejects the claim. Alternatively, the algorithm can return a confidence measurement of the claim's validity. Verification applications include those that authenticate identity during point-of-sale transactions or that control access to computers or secure buildings.

Performance statistics for verification applications differ substantially from those for identification systems. The main performance measure for identification systems is the system's ability to identify a biometric signature's owner. More specifically, the performance measure equals the percentage of queries in which the correct answer can be found in the top few matches.

For example, law enforcement officers often use an electronic mug book to identify a suspect. The input to an electronic mug book is a mug shot of a suspect, and the output is a list of the top matches. Officers may be willing to examine only the top twenty matches. For such an application, the important performance measure is the percentage of queries in which the correct answer resides in the top twenty matches.

The performance of a verification system, on the other hand, is traditionally characterized by two error statistics: false-reject rate and false-alarm rate. These error rates come in pairs; for each false-reject rate there is a corresponding false alarm. A false reject occurs when a system rejects a valid identity; a false alarm occurs when a system incorrectly accepts an identity.

In a perfect biometric system, both error rates would be zero. Unfortunately, biometric systems aren't perfect, so you must determine what trade-offs

Biometric Organizations

Kirk L. Kroeker, Computer

Although poised for substantial growth as the marketplace begins to accept biometrics, recent events have demonstrated that the fledgling industry's growth could be severely constricted by misinformation and a lack of public awareness.

In particular, concerns about privacy can lead to ill-informed regulations that unreasonably restrict biometrics use. The lack of common and clearly articulated industry positions on issues such as safety, privacy, and standards further increase odds that governments will react inappropriately to uninformed and even unfounded assertions regarding biometric technology's function and use.

Two organizations, the International Biometric Industry Association and the Biometric Consortium, aim to improve this situation.

International Biometric Industry Association

A Washington, D.C.-based trade association, the IBIA seeks to give the young industry a seat at the table in the growing public debate on the use of biometric technology. The IBIA focuses on educating lawmakers and regulators about how biometrics can help deter identity theft and increase personal security.

In addition to helping provide a lobbying voice for biometric companies, the IBIA's board of directors has taken steps to establish a strong code of ethics for its members. In addition to certifying that the consortium will adhere to standards for product performance, each member must recognize the protection of personal privacy as a fundamental obligation of the biometric industry.

Besides promoting a position on member ethics, the IBIA recommends

- safeguards to ensure that biometric data is not misused to compromise any information;
- policies that clearly set forth how biometric data will be collected, stored, accessed, and used;
- limited conditions under which agencies of national security and law enforcement may acquire, access, store, and use biometric data; and
- controls to protect the confidentiality and integrity of databases containing biometric data.

The IBIA is open to biometric manufacturers, integrators, and end users (http://www.ibia.org).

Biometric Consortium

On 7 December 1995, the Facilities Protection Committee (a committee of the Security Policy Board established by US President Bill Clinton) chartered the Biometric Consortium. With more than 500 members from government, industry, and academia, the BC serves as one of the US government's focal points for research, development, testing, evaluation, and application of biometric-based systems. More than 60 different federal agencies and members from 80 other organizations participate in the BC.

The BC cosponsors several biometric-related projects, including some of the activities at NIST's Information Technology Laboratory and work at the National Biometric Test Center at San Jose State University. The BC also cosponsors NIST's Biometrics and Smart Cards laboratory, which addresses a wide range of issues related to the interoperability, evaluation, and standardization of biometric technologies and smart cards, especially for authentication applications like e-commerce and enterprise-wide network access.

In September 1999, the BC held its annual conference on the convergence of technologies for the next century. The conference highlighted and explored new applications in e-commerce, network security, wireless communications, and health services. It also addressed convergence of biometrics and related technologies like smart cards and digital signatures.

The BC's Web site and its open listserv are two of the consortium's richest resources (http://www.biometrics.org).

Kirk L. Kroeker is associate editor at Computer *magazine. Contact him at kkroeker@computer.org.*

you're willing to make. If you deny access to everyone, the false-reject rate will be one and the false-alarm rate will be zero. At the other extreme, if you grant everyone access, the false-reject rate will be zero and the false-alarm rate will be one.

Clearly, systems operate between the two extremes. For most applications, you adjust a system parameter to achieve a desired false-alarm rate, which results in a corresponding false-reject rate. The parameter setting depends on the application. For a bank's ATM, where the overriding concern may be to avoid irritating legitimate customers, the false-reject rate will be set low at the expense of the false-alarm rate. On the other hand, for systems that provide access to a secure area, the false-alarm rate will be the overriding concern.

Because system parameters can be adjusted to achieve different false-alarm rates, it often becomes difficult to compare systems that provide performance measurements based on different false-alarm rates.[1,2]

EVALUATION PROTOCOLS

An evaluation protocol determines how you test a system, select the data, and measure the performance. Successful evaluations are administered by independent groups and tested on biometric signatures not previously seen by a system. If you don't test with previously unseen biometric signatures, you're only testing the ability to tune a system to a particular data set.

For an evaluation to be accepted by the biometric community, the details of the evaluation procedure must be published along with the evaluation protocol, testing procedures, performance results, and representative examples of the data set. Also, the information on the evaluation and data should be sufficiently detailed so that users, developers, and vendors can repeat the evaluation.

The evaluation itself should not be too hard or too easy. If the evaluation is too easy, performance scores will be near 100 percent, which makes distinguishing between systems nearly impossible. If the evaluation is too hard, the test will be beyond the ability of existing biometric techniques. In both cases, the results will fail to produce an accurate assessment of existing capabilities.

An evaluation is just right when it spreads the performance scores over a range that lets you distinguish among existing approaches and technologies. From the spread in the results, the best performers can be determined along with the strengths and weaknesses of the technology. The strengths and weaknesses

Practical Systems for Personal Fingerprint Authentication

Lawrence O'Gorman, Veridicom Inc.

Before the mid-1990s, optical fingerprint-capture devices were bulky (about the size of half a loaf of bread) and expensive (costing anywhere from $1,000 to $2,000). Technological advances have brought the size and cost down dramatically; the new solid-state sensors cost less than $100 and occupy the surface area of a postage stamp. Previously used primarily for government applications, fingerprint authentication technology is now steadily progressing into the private sector for the many applications requiring both convenience and security.

The small size and cost of these devices can provide secure access to desktop PCs, laptops (as shown in Figure A), the Web, and most recently, to mobile phones and palm computers. Automobile manufacturers are building prototype cars with access and personalization (of seat position, radio channels, and so on) that are controlled by fingerprint authentication devices. Someday soon, when the sensor is small, inexpensive, and low power enough to build into a key fob, many of us will carry a *universal key* to facilitate secure access to everything from front doors to car doors, computers, and bank machines.

Figure A. Fingerprint authentication devices will find increasing application in securing laptops. The fingerprint sensor is the small rectangle to the bottom right of the keyboard.

Fingerprint sensors

The companies developing this technology have used different means for fingerprint capture, including electrical, thermal, or other means. For example, a capacitive-sensing chip measures the varying electrical-field strength between the ridges and valleys of a fingerprint, as shown in Figure B. A thermal sensor measures temperature differences in a finger swipe, the friction of the ridges generating more heat than the nontouching valleys as they slide along the chip surface. Some companies are working on optical and hybrid optical/electrical capture devices whose optics have shrunk to about 1.5 cubic inches.

Portable computing

One of the first widespread applications of personal authentication will be for portable computing. In terms of financial losses for corporate computing, laptop theft in 1999 ranked third at $13 million

detected during the evaluation indicate which applications the technology can address adequately.

Technology

The most general type of evaluation tests the technology itself. You usually perform this kind of evaluation on laboratory or prototype algorithms to measure the state of the art, to determine technological progress, and to identify the most promising approaches. This evaluation class includes the Feret (face recognition technology) series of face recognition evaluations and the National Institute of Standards and Technology (NIST) speaker recognition evaluations.

The best technology evaluations are open competitions conducted by independent groups. In these evaluations, test participants familiarize themselves with a database of biometric signatures in advance of the test. They then test algorithms on a sequestered portion of the database. This practice allows systems to be tested on data that the participants haven't seen previously. The use of test sets allows the exact same test to be given to all participants.

Evaluations typically move from the general to the specific. The first step is to decide which scenarios or applications need to be evaluated. Once the evaluators determine the scenarios, they decide upon the performance measures, design the evaluation protocol, and then collect the data.

Scenario and operational

Scenario evaluations measure overall system performance for a prototype scenario that models an application domain. An example is face recognition systems that verify the identity of a person entering a secure room. The primary purpose of this evaluation type is to determine whether a biometric technology is sufficiently mature to meet performance requirements for a class of applications. Scenario evaluations test complete biometric systems under conditions that model real-world applications. Because each system has its own data acquisition sensor, each system is tested with slightly different data. One scenario evaluation objective is to test combinations of sensors and algorithms. Creating a well-designed test, which evaluates systems under the same conditions, requires that you collect biometric data as closely as possible in time.

To compensate for small differences in biometric signature readings taken over a given period, you can use multiple queries per person. Because scenario eval-

behind financial fraud ($39 million) and theft of proprietary information ($42 million). However, the problem goes far beyond loss of the computer; compromised information security may incur far greater business cost.

Furthermore, laptops frequently provide access to a corporate network via software connections (complete with stored passwords on the laptop). The solid-state fingerprint sensor—small, inexpensive, and low power—solves these problems. With appropriate software, this device authenticates the four entries to laptop contents: login, screen-saver, boot-up, and file decryption.

Cryptography

Personal authentication also can come into play in cryptography, in the form of a private-key lockbox, which provides access to a private key only to the true private-key owner via his fingerprint. The owner can then use his private key to encrypt information relayed on private networks and the Internet. Although good encryption methods are very difficult to break, the Achilles heel in many encryption schemes is ensuring secure storage of the encryption key (or private key). Frequently, a 128-bit or higher key is safeguarded only by a 6-character (48-bit) password. A fingerprint provides much better security and—unlike a password—is never forgotten. In the same way, a fingerprint-secured lockbox can contain digital certificates or more secure passwords—ones that are much longer and more random than those commonly chosen—for safeguarding e-commerce and other Internet transactions. These schemes assure a user both security of electronic transactions as well as personal privacy.

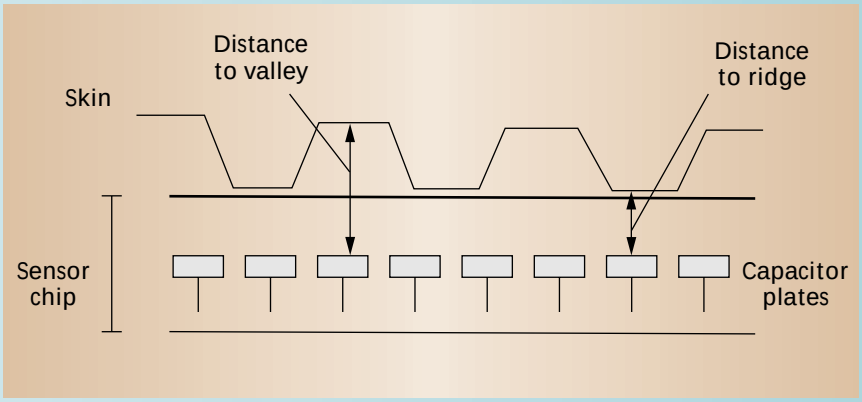


Figure B. Capacitive sensing is one way devices distinguish between fingerprint patterns. Fingerprint ridges and valleys touch the sensor's surface. The sensor measures the distances to the skin to capture an image of the fingerprint.

uations test complete systems under field conditions, they cannot be repeated. You can only attempt to retest under similar conditions.

An operational evaluation is similar to a scenario evaluation. While a scenario test evaluates a class of applications, an operational test measures performance for a specific algorithm for a specific application. For example, an operational test would measure the performance of system X on verifying the identity of people as they enter secure building Y. The primary goal of an operational evaluation is to determine if a biometric system meets the requirements of a specific application.

FACE RECOGNITION

Although you can choose from several general strategies for evaluating biometric systems, each type of biometric has its own unique properties. This uniqueness means that each biometric must be addressed individually when interpreting test results and selecting an appropriate biometric for a particular application.

In the 1990s, automatic-face-recognition technology moved from the laboratory to the commercial world largely because of the rapid development of the technology, and now many applications use face recognition.[3] These applications include everything from controlling access to secure areas to verifying the identity on a passport. The most recent major evaluations of this technology took place between September 1996 and March 1997 with the Feret.[4,5] The Feret tests were technology evaluations of emerging approaches to face recognition. Research groups were given a set of facial images to develop and improve their systems. These groups were tested on a sequestered set of images, which required the participants' systems to process 3,816 images.

The Feret evaluation measured performance for both identification and verification, and provided performance statistics for different image categories. The first category consisted of images taken on the same day under the same incandescent lighting. This category represented a scenario with the potential for achieving the best possible performance with face recognition algorithms. Each of the following three categories became progressively more difficult, with the final category consisting of images taken at least a year and a half apart.

Table 1 summarizes the verification performance results for the best algorithms in each category. The results are from a database of 1,196 people. The results in Table 1 show that illumination and time between acquisition of each image can significantly affect face recognition performance.

Automotive

A third application is for automobiles. A sensor, located either in the car door handle or in a key fob, could unlock the car, and another in the dashboard could control the ignition. Reliability is a concern, however, because automobile sensors must function under extreme weather conditions on the car door and high temperature in the passenger compartment. And a key fob sensor must be scratch-, impact-, and spill-resistant. It also must be able to sustain an electrostatic discharge of greater than 25 kV—no small dose of voltage for a chip.

Despite these concerns, automotive parts manufacturers are forging ahead. Safeguards, such as protecting the sensor within an enclosure or placing it in a protected location on the car, are under consideration.

Pioneers in practical fingerprint authentication

Recognizing the potential of small and inexpensive fingerprint sensors, several companies have developed technologies for this purpose. Among these are the following:

- **Authentec** (http://www.authentec.com) makes FingerLoc, a biometric identification subsystem. It uses CMOS and electric-field imaging.
- **Veridicom** (http://www.veridicom.com), **STMicroelectronics** (http://us.st.com), and **Infineon** (http://www.infineon.com) all have products that use CMOS and capacitive imaging (5thSense, TouchChip, and FingerTIP, respectively).
- **Thomson-CSF** (http://www.tcs.thomson-csf.com) has developed FingerChip, which also uses CMOS, but utilizes thermal imaging.
- **Who?Vision's TactileSense** (http://www.whovision.com) images via an optoelectrical polymer mounted on a thin-film transistor.
- **Identix** (http://www.identix.com) makes optical fingerprint readers.

The small size and low cost of these new fingerprint sensors make them an ideal human interface to secure systems. These and many more applications will soon incorporate personal biometric authentication. If the current trends continue, the public sector can expect to see such devices increasingly incorporated into everyday life. ❖

Reference

1. *CSI/FBI Computer Crime and Security Survey*, Computer Security Institute, San Francisco, 1999.

Lawrence O'Gorman is chief scientist for Veridicom Inc. His research interests include image processing and pattern recognition. O'Gorman has a PhD from Carnegie Mellon University, an MS from the University of Washington, and a BASc from the University of Ottawa, all in electrical engineering. He is a Fellow of the IEEE and of the International Association for Pattern Recognition. Contact him at log@veridicom.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.