A review of biometric experts' evaluations of the biometric modalities most likely to effectively combat identity fraud

Galaxy Samson Edo
Capella University

ABSTRACT

This article reviews biometric experts' evaluations of the biometric modalities most likely to effectively combat identity fraud. The expert views on the viability of biometrics are expressed through formal evaluation of biometric systems, comparisons of one type of biometrics to another and also comparison between unimodal and multimodal biometrics. The review briefly surveyed the current state of identity fraud prevention efforts, based mostly on improving organizational or public attention to the issue, including the introduction of red flags. Overall, the review found that experts prefer fingerprint biometrics to all other unimodal biometrics, but that they are also shifting their support to multimodal systems

Keywords: biometrics, biometric experts, biometric evaluation, biometric preference, fingerprinting, iris identification, keystroke, DNA, voice, identify fraud, identity theft, unimodal and multimodal.

INTRODUCTION

Identity fraud is the criminal use of false identities or fraudulent identification documents (Wilcox & Regan, 2002). This criminal event occurs when one person takes identifying information belonging to somebody else and uses such identifying information for abuse, without authority (Wilcox & Regan, 2002). Identity fraud is predominantly means of perpetrating a financial fraud, related to bank fraud and credit-card fraud. However, there are many other types of identity fraud related to computer and telecommunications, access-device fraud, tax-refund fraud, social-program fraud, mail fraud, and terrorism (Wilcox & Regan, 2002). Since 9/11, identity fraud has become "a growing national concern" in "its potential national security implications" (Pinheiro, 2004, Para. 2).While biometrics is an automated use of unique human physiological or behavioral characteristics to determine or verify a person's unique identity (Kleist, 2007).

Gordon and Wilcox (2003) claimed, "Identity fraud is a national and global threat to the security of nations and their citizens, the economy, and global commerce, as it facilitates a wide range of crimes and terrorism" (p. 4). The examination of the aftermath of the 9/11 terrorist act brought to light the extent to which the use of fraudulent identification is not only a significant component of fraud but also of terrorism. The recent debacle over a Yemeni woman whose identity was stolen to mail a package containing explosives to the United States also attests that identity theft is indeed a global threat (Dozier, 2010).

According to Stana (2002) "the events of September 11, 2001, have heightened concerns about the contributory role that identity fraud plays in facilitating terrorism and other serious crimes" (p. 1). The creation and use of a false identity from fraudulent documents or stolen identity in the commission of a crime has long been used by criminals and criminal organizations to facilitate criminal activities and avoid detection (Wilcox & Regan, 2002).

In 2007, the Federal Trade Commission received over 800,000 consumer fraud and identity complaints, representing an increase of 21% from the year before. Jarillo, Pedrycz and Reformat (2008) estimated that identity theft had resulted in $21 billion lost during 2003. Identity theft was estimated in 2006 to have cost the economy $49 billion per year. At present estimate, identity fraud costs U.S. businesses $53 billion annually (Swartz, 2009) and 3.7% of U.S. adults were victims of identity fraud in 2006 (Goldwasser & Anderson, 2007). Gregory (2008) also noted that by one measure ID fraud cases have risen from 9,000 in 1999 to 77,500 in 2007. Moreover, this number was expected to double in the next five years. Online shopping is one of the engines of this escalation, as a good deal of fraud occurs in online shopping contexts in the form of phishing, in which the fraudster acquires usernames, passwords or credit card numbers (Gregory, 2008). Studies of Facebook users have also found that one in four users inadvertently exposed personal data to strangers. Banks seem to be particularly vulnerable. All of the ways in which banks in particular are working to protect consumers from identity fraud are called cyber armor by some (Goldwasser & Anderson, 2007).

The Identity Fraud Safety Scorecard is calculated each year by a research company that monitors identity fraud prevention in the credit card business. The scorecard was established to heighten business awareness of the extent of identity fraud losses, which, according to McCollum (2005) added up to $52.6 billion, affecting 9.3 million people in 2008. At present, most companies seek to protect customers by placinglimits on types of transactions. But McCollum found that the fact that two-thirds of cardholders are asked to enter their Social Security Number to access account information was a glaring finding as was the fact that few

cardholders had mechanisms in place to alert users to unusual account activity. The elimination of the SSN as part of the identity process for access to most cardholders account would, in any case, limit the damage due to security breach and identity fraud. Still, the survey of cardholders found no regular use of biometric elements as a way to improve security.

Mercuri (2006), however, reported on another study that suggested that estimates of identity fraud may be exaggerated. In a report by the Federal Trade Commission, it was found that 85% of identity fraud involved the misuse of existing accounts, 17% involved the opening or new accounts and only 17% entailed the misuse of personal information. Such activities were only criminalized in 1998 through the Identity Theft and Assumption Deterrence Act that made using a false identity to use an account illegal. With regard to fraud in financial transactions, businesses carry 90% of the losses while consumers only shoulder 10% of these frauds.

The average loss to an individual from identity theft of account holdings is roughly $1100 per person. Consumers are in fact now buying insurance policies to protect them from the negative consequences of fraud. Also, studies have found that while computer-based identity fraud has risen, six times more identity crimes begin with a stolen wallet. Online, 20% of fraud cases resulting from phishing attacks, with consumers responding to illegitimate websites. In explaining why, with rising online or computer fraud, companies have not acted more forcefully, Mercuri (2006) suggested that the fact that most losses can be tax write-offs perhaps reduces company motivation to do so. This may also be why so many companies continue to rely upon simple rubrics for identity numbers. Most companies do not suffer enough loss from identity fraud to make it a high priority issue for them.

BIOMETRIC EXPERTS' EVALUATIONS

Expert views on the viability of biometrics are expressed through formal evaluation of biometric systems, comparisons of one type of biometrics to another and also comparison between unimodal and multimodal biometrics. Experts conduct evaluations based on established or emerging evaluative models in order to determine best practice in biometrics (Coventry, 2005; Dekking & Hensbergen, 2009; Gorodnichy, 2009; Jain & Pankanti, 2001; Mane & Jadhav, 2009; Schuckers, 2003; Sulovska & Adamek, 2010; Volner & Bores, 2009; Wechsler, 2010).

Hong et al. (2005) pointed out that one of the problems with biometric security systems is that they are dynamic, due to changes in the biometric measures on persons over time, and these changes are difficult to control. As a result, more effort is needed to evaluate the effectiveness of a biometric system across time. Criteria for evaluating a biometric feature would include its universality, uniqueness, permanence and collectability. Biometric features can also be evaluated according to their performance (or accuracy), acceptability (or how much people accept it) and circumvention (or how easy it is to cheat the system). At present, no biometric features satisfied all of these criteria, causing Hong et al. to favor multimodal systems, which may use multiple sensors, biometrics or units or multiple instances or impressions of the same biometrics. A database could also accept multiple representations and matching algorithms from the same input biometric, combined with different approaches to feature extraction.

As biometrics increased in use, standardizations have also developed, as the industry approaches maturity. Thus, the data must be interoperable and interchangeable according to certain standards. When it comes to an expert evaluating the performance of a biometrics system, they look at the suitability of the biometric used according to its universality, uniqueness, permanence and security (Hong et al., 2005). Mansfield and Wayman (2002)

agreed that an evaluation would consider the strengths and weaknesses of the biometric according to standards. The results of an evaluation were often quantified by representation of the receiver operating characteristic (ROC) or the Detection Error trade-off curve (DET).

Hong et al. (2005) tested an evaluation for a fingerprint-based biometric system. Since fingerprints are easily corrupted or damaged, image quality and feature quality checks are a basic part of an evaluation and analysis. Many fingerprint evaluations only measure algorithm performance, leading Hong et al. to find them of limited quality. Overall, however, evaluations are undertaken to determine how accurate the biometric is, and to uncover weaknesses in the system and fix them.

Biometrics also needs to be evaluated more often than conventional methods because they employ complex pattern recognition modules that can commit two types of errors, a false accept and a false reject (Podio & Dunn, n.d.). As a result, biometrics presents researchers with several additional evaluative issues. A biometric can have errors at the sensor point, called information limited behavior, errors at the feature extraction level, resulting in representation limited behavior, and with a modeler/matcher which identifies the invariant elements of the input, called invariance limited behavior (Wayman et al., 2005). These limitations affect the receiver operating characteristics (ROC) of the system as a whole, which measures the ratio of the above rates, a basic performance measure of a biometric system. Jain and Pankanti (2001) demonstrated in detail how well the feature extractor and matcher work in a fingerprint-based biometric system.

Jain and Pankanti (2001) also argued that at present some of the fundamental questions about the accuracy of a practical biometric system, involving the inherent discriminant information in the input signal, have "only been answered in a very limited way for most biometrics modalities" (p. 6). Most systems fail at present to be able to discriminate between the fingerprint of the same person before and after an accident, for example, degrading the fingerprint quality. A recent study found that due to residue from previous fingerprints, 4% of current fingerprint images were not useful for personal authentication. Jain and Pankanti (2001) also noted that in many cases biometrics have yet to develop sophisticated frameworks for representation and feature extraction, still relying on rather crude verification data (as in the case of speaker variation research, which continues to be approximate). At present, Jain and Pankanti (2001) argued that biometrics evaluation also fails to provide experts with the capability of predicting how a biometric device will perform in the real world.

Testing types of biometrics.

Tappert, Cha, Villani, and Zack (2010) tested a novel keystroke biometric system for long-test input in more than 100 experiments with participants using either copy or free text on desktops or laptop computers. Keystroke biometrics measures the typing characteristics of the user, believed to be unique to an individual and extremely difficult to duplicate. Keystroke identity is established by almost unconscious rhythms in keystroke pressing, measuring in terms of key press times, key release times, keystroke duration times and keystroke transition times. Most of the research on keystroke biometrics thus far has been experimental, without consideration of the public. And yet keystroke biometrics is particularly appealing to the public because they are not asked to participate in an activity that they are unfamiliar with. Keystroke security is also inexpensive and keystrokes can be subjected to dynamic verification. Not only can keystroke data be sent over the internet but also if based on a long-test analysis then

powerful statistical measurements can be applied to it, greatly enhancing its security (Tappert et al., 2010).

Overall, "the published literature is optimistic about the potential of keystroke dynamics to benefit computer system security" (Tappert et al., 2010, p. 32). Here as in other studies the ROC curve (representing the trade-off between false accept rate and false reject rate) was used to describe the performance of the system. The system was found to be accurate in identifying a user so long as the same type of keyboard was used to produce the input. Longitudinal experiments also found that keystroke identification continued to identify users for years. It was also found that the input of 300 keystrokes was sufficient to identity an individual. Keystroke analysis was also found to be helpful in online test-taking and email situations. While the ROC curves provided the possibility of a number of possible tradeoffs between FAR and FRR, Tappert et al. (2010) concurred with other researchers that the ROC curve should be used to evaluate systems exclusively, as it does not reflect a number of other important variables.

Dekking and Hensbergen (2009) discussed problems with the evaluation of an iris recognition system, described as one of the most promising biometric technologies. A mathematical code called the Hamming Distance is the most common method used to evaluate the quality of matching between input and template of iris biometric data. In reviewing the mathematics underlying the Hamming Distance, Dekking and Hensbergen determined that some of the parameters made use of in the mathematics are based on unfulfilled assumptions. As a result, most evaluations of the iris identification system are overly optimistic in assessing their reliability. Dekking and Hensbergen simply intended to demonstrate that when experts evaluate the effectiveness of biometric systems for combating identity fraud, the tools they make use of to do so can also have problems.

Expert preferences.

Nonetheless, in practical terms, Jain and Pankanti (2008) appeared, as experts in the field, to continue to favor fingerprint, face and iris recognition biometrics. Some of the advantages of fingerprinting are that the sensors for capturing prints are cheap and small enough to be embedded in consumer products, even though these small sensors have high error rates. Face recognition also has the potential to become a possible mainstream biometric usage because it exploits the capability of many electronics that now have built-in cameras. Face recognition was also accurate in controlled settings, but developed problems in less controlled settings with changes of pose, lighting, expression, and facial accessories interfere. Face recognition is also very unreliable when captured by video cameras, "in which subjects do not present themselves in front of the camera in predetermined poses" (Jain & Pankanti, 2008, p. 2). Iris identification was described by Jain and Pankanti (2008) as accurate and swift, which is why it was adopted by the UK customs in its Iris Recognition Immigration System. That said, the random patterns in the iris are so complex that no known human experts, with the naked eye, are capable of determining whether two iris images match, that is, only the machine not man can detect variation, meaning that iris matches are not usable in courts of law. Jain and Pankanti (2008) again addressed the fact that biometric systems make decisions based on imperfect matches, meaning that they can generate two types of errors, the false accept and the false reject.

The general rubric of acceptability for a biometric system is a false reading rate of one mistake in every one thousand assertions of a match (Jain & Pankanti, 2008). Recent tests of biometric systems have found error rates higher than this acceptable limit. Higher quality

images and refinement of feature extraction are required to reduce false matches. Also, biometric systems are prone to being fooled by spoof traits, or presentation of plastic copies of hands or fingers (Jain & Pankanti, 2008). Thus, in addition to the biometric of choice in any system, Jain and Pankanti (2008) argued that all biometric systems need to develop additional sensors to detect body heat and other signs of life to prevent acceptance of spoof images. They also favored multimodal biometrics, as different traits or multiple instances of traits can provide "irrefutable proof of legitimate identity" (p. 2).

Coventry (2005) argued that biometric technology, insofar as it is based on the features of a person, and not on a token or a document used for identification purposes, has the potential to greatly reduce authentication and identity fraud. Biometrics emerged as a commercially available technology thirty years ago, but has witnessed a surge of interest in the current security climate. At the same time, Coventry regretted that most biometrics methods developed thus far are system-centered and was developed without input from the usability community, or experts on technology usability and acceptance by the public.

Coventry (2005) argued that more research is required into the usability aspects of biometrics in order to improve implementations. Biometrics are used both for identifying the person and then for verifying the identification by matching it against a database template. Thus far only fingerprinting and retinal and iris scanning have been found to be able to accurately identify a person in a large database. Facial systems, however, because they must be scrutinized at length by a human observer, may not be suitable for real-time identification. Indeed, the process time required for much biometric identification makes its use for real-time ID less likely.

Fingerprints, in terms of usability, also have problems, with a great variety in quality of prints based on gender and socioeconomic status. Optical systems are also negatively affected by "dirt, cold fingers and finger damage and are otherwise prone to fraud" (Coventry, 2005, p. 185). Users also have concerns about the use of fingerprinting, such as hygiene, though the traditional stigma of criminality attached to their use has diminished. Context is also critical. Thus far, commercial uses of biometrics involve controlling physical access, in places ranging from secure locations such as prisons to company premises. In this context, biometrics in the form of hand geometry and fingerprinting are most commonly used (Coventry, 2005).

Biometrics is also increasingly used for immigration and border control purposes. Iris, fingerprint and facial biometrics have begun to be incorporated into passports (Ashbourne, 2004). Distance biometrics has also emerged, involving screening from video cameras in various locations. Biometrics is also being used for ID cards or driver's licenses and to prevent welfare fraud. Biometrics can also ensure persons making transactions on the Internet that it is authenticated. Coventry (2005) specifically studied the use of biometrics in the context of bank ATM machines to address the degree to which public usability issues can compromise effectiveness. He argued that in the ATM context an only contact silicon device for thermal swipes is feasible, with optical sensors being impractical because of dirt accumulation. He also argued, however, that fingerprints given through being previously embedded on smart cards is probably the most practical approach to biometrics in an ATM context. Facial biometrics would be more difficult, because of height differences, the requirement that users stand in a specific spot and assume a neutral expression, the fact that light must be uniform, and that there is a high chance of false negatives.

Coventry (2005), pursuing all possibilities of fraud in user contexts, even expressed concern that people could attempt to defraud the system through "decapitation of a legitimate user or the use of masks" (p. 188). Iris biometrics are accurate but at present require the user to

put their eye close to the sensor, which the public might find difficult to achieve, and these systems are also very expensive. Retina biometrics is ruled out because the technique is invasive and requires training to use. Hand biometrics has similar problems as face recognition as different hand sizes and accumulation of dirt would lead to too many false readings. Signature use is attractive to Coventry for the ATM because of the long association in the public's mind between financial authorization and giving signatures.

To make the decision as to which biometric method is most usable in a specific context, Coventry (2005) argued that evaluation methods of biometrics must be improved. Though a number of private and public testing laboratories have been set up to undertake biometric evaluation, as yet a standard for biometrics evaluation has not been formally established. This is because, while theoretical grounding for some techniques is impressive, how they actually perform in the real-world is another, less studied matter. As noted previously, most performance evaluation of biometrics is based on false accept rates and false reject rates. False accept rates involve the wrong persons being able to access the system, while false reject rates involve legitimate persons being denied access to the system (Coventry, 2005). The two factors are interconnected, as when one improves the other worsens. The methods by which these figures are established, however, are to Coventry untested.

Moreover, performance estimates can underestimate real world performance. Many biometric systems often fail to live up to expectations because they "prove unable to cope with the enormous variations among large populations, or fail to take into account people's needs and behaviors" (Coventry, 2005, p. 193). For many biometric systems, the false reject rate increases for persons over 50, for reasons that are as yet unclear. Other factors that must be considered in evaluating biometrics are the failure to enroll rate, which identifies people who can never use the system, and the failure to acquire rate, or the number of users unable to generate an image when using the device (Coventry, 2005). As a result of these problems, a good biometric system needs a fallback strategy that allows these users another way into the system. All of these factors then must be factored into an evaluation of biometrics, that is, the user base influences performance, and the resulting usability factor is a major issue (Coventry, 2005).

Focus groups were recommended by Coventry (2005) in order for designers to develop a better understanding of user behavior and limitations. Field trials are also imperative. Issues involving enrollment in the system, that is, submitting one's information to a database, which itself necessitates a high quality image to create an accurate template, must also be addressed. To generate a high quality enrollment image users must be educated about the biometrics of fingerprints (for example, how to place the finger on the sensor to gain an optimal image), trained in how to use the technology, learn how the software interface will support them, be trained in the steps involved in the interaction, and be given time to explore and learn how to use the system in what is termed "playtime" (Coventry, 2005, p. 196). All of these factors must be taken into consideration. Unfortunately, too many designers of biometrics assume that they are easy to use, and then have no answers for consumers challenged by the system. Thus, "the user's interaction with the biometrics device and the feedback provided by the system are crucial for success" (Coventry, 2005, p. 197).

Still another problem in everyday use of biometrics is that some people will be rejected and unable to use the system, meaning that a means of bringing these outliers back into the system must be devised (Coventry, 2005). This entails having an exception handling method for dealing with the injured, the sick, the elderly and others often rejected by biometric systems because they cannot generate a clean image.

Overall, then, Coventry (2005), using some of the elements of the technology acceptance model, found that the technology must be socially acceptable, appropriate for the given context, fulfill a perceived need, be understandable, usable and not compromise one's privacy. With regard to perceived need, up until recently few members of the public saw the need for such security. Since 9/11, however, security concerns have moved to the front of consumer concerns and thus biometrics would now find an accepting audience (Stana, 2002). In a survey of response to security measures, 75% of respondents reported that biometrics is more secure than traditional security methods. At the same time, people are wary of being rejected by such futuristic technology, and rejection itself can generate emotions that would invalidate efforts to resubmit data (Coventry, 2005). Of course, the major consumer concern is privacy, about the capacity for such devices to expose a genetic disorder or HIV or otherwise invade their privacy. One survey found that the public was not so much wary of biometrics per se, but of the dangers involving third parties who could intercept data and use it for applications not permitted by the consumer (Hoonakker, Bornoe, & Carayon, 2009). Finally, Coventry felt that while biometrics has established an impressive research base, its usability in response to public issues still remains a question. He was also not entirely convinced that the public would accept biometrics. The primary takeaway point from Coventry's analysis, however, was that there remains a gap between theory and practice in biometrics. Moreover, while scientists devise theoretical advances of technology, experts on the use biometrics for combating identity fraud make their evaluations based on the actual usability of the technology and the likelihood of public acceptance (using the Technology Acceptance Model as a base). Thus, the experts evaluate theoretical developments in light of application to identity theft problems.

Jain et al. (2000) described a biometric system as pattern recognition with an enrollment module and an identification module. In evaluating the effectiveness of the system each area must be assessed for performance quality, acceptability to the public and circumvention, or how easy it is to defraud. Biometric systems also use a verification/authentication mode or a recognition/identification mode. Overall, the system was evaluated according to accuracy, speed and storage. Again, a system can fail either by accepting an impostor as valid, in a false match, or rejecting a valid individual, a false nonmatch, and evaluation can calculate the false match and nonmatch rates in order to assess its security quality. The receiver operating characteristics (ROC) graph measures the trade-off between FMR and FNR, and is believed to be a comprehensive measure of system accuracy. High-security biometric systems operate with a small FMR acceptable, forensic applications favor catching criminals and so operate matchers at a high FMR rate, while civilian applications try to operate with both a low FNR and low FMR. Where FMR meets FNR is called the equal error rate, and "may often be used as a terse descriptor of system accuracy" (Jain et al., 2000, p. 95). In evaluating the accuracy performance of a type of biometrics, the system is considered acceptable if the risks at a given ROC point are acceptable and unacceptable if the risks are not. Acceptability also often measures the amount of time it takes to record the biometric input, with ATM banking contexts for example requiring real-time results, but forensic applications not. The cost of the system is also considered. In reviewing the various types of biometrics with this rubric, then, Jain et al. found face recognition to be difficult and often unreliable.

A facial thermogram, by contrast, charts out through heat sensors the unique facial signature that occurs when heat passes through the facial tissue. It is claimed that facial thermograms are unique to individuals, and cannot be altered, because they represent the flow of blood through the veins, by plastic surgery (Jain et al., 2000). The fact that a face thermogram

can be taken in ambient light situations and remains invariant to any change of expression, makes it more accurate. Fingerprinting is favored by Jain et al. (2000), but it remained that the public has suspicions about fingerprinting, and in some populations fingerprints are likely to be degraded in quality. Hand geometry is also easy to use and cheap, but for Jain et al. has a low level of discriminative capability. Retinal scanners are highly accurate, but also require a good deal of public cooperation in delivering the image and expensive to use.

Thus, retinal scanning is only used in high security areas. Irises have also been found to be unique, even between identical twins, and current methods requiring a high level of public cooperation are being replaced by more user-friendly methods. Jain et al. (2000) also liked signature use because the public has long ago accepted giving signatures to participate in any number of commercial endeavors. Speech patterns are also unique but for Jain et al. provided information with too much variance. Speech systems are also sensitive to background noise and other disturbances. Though only evaluating the relative effectiveness of different kinds of biometric data, it appeared that Jain et al. favored development of hand geometry and iris identification, but generally believed that they will only replace fingerprinting and signature writing when their recording becomes more discriminate and less invasive.

Schuckers (2003) made use of a beta-binomial distribution to evaluate the matching performance of biometric identification devices. Again, the distribution was made use of to assess the variability in the false match and false non-match rates of a biometric device testing a number of different subjects. Again, these rates are recorded in the Receiver Operating Characteristics, the most common method in use for evaluating the overall performance of a BID. The ROC plots a curve of false non-match rates against false match rates. But as yet no consensus has emerged on how to assess the performance of a biometric device when two or more individuals are tested (Schuckers, 2003).

Schuckers (2003) argued that a beta-binomial can do this, where a binomial could not. The purpose of the evaluation was to take extraneous variability factors into better consideration and cerate confidence intervals. After detailing the elements of the device, examples are given to illustrate its use to estimate overall system matching performance. Schuckers concluded on the basis of tests “the Beta-binomial seems to be an appropriate distribution for assessing the matching performance of a BID” (p. 529). In this way, using techniques of this kind, experts evaluate the effectiveness of various biometric methods, to determine which method results in the most accurate findings.

Mitra, Savvides, and Brockwell (2007) provided still another way to authenticate a biometric authentication system, using statistical methods. Again, they argued that biometric systems will only become acceptable and stakeholders change from the use of passwords if authentication systems can prove its advantages. The core of biometric evaluation is determining the extent to which the biometric samples obtained from persons submitted to the system match with templates stored in the database, and usually synthesized from training data. Mitra et al. (2007) acknowledged the importance of the receiver operating characteristic as a diagnostic device in authentication of biometrics using matching performance methods. Indeed, ROC curves are the most common evaluation criterion made use of in current evaluation studies of biometric systems.

But Mitra et al. (2007) considered the problem of watch lists, and how or if biometric evaluation could zero in on the system performance in detecting persons on a watch list. A watch-list consists of a database within a database of persons who are of some interest, and whose names, for various security-related systems, have been placed on a watch list, such as a do

not fly list at airports. The major problem with most watch list systems is that they have a tendency to produce too many false alarms. Watch lists are known to be behind the stopping and questioning at airports of persons who by all other appearances or measures pose no threat. Most watch lists perform so badly, Mitra et al. argued, because they are name-based as opposed to biometrics-based systems. Some biometric systems have been used to develop or monitor watch lists with one study finding that the Face Recognition Test did well in matching faces with persons on watch lists. At the same time, it has been generally found that the watch list matching ratio gets poorer as the size of the watch list grows. Thus, Mitra et al. felt that it was necessary to develop a performance evaluation of a biometric system using a statistical framework that predicts misclassification rates and false alarms on watch lists.

The statistical method was developed to account for the increasing size of the database as a factor in considering whether or not too many false alarms are being detected. Thus, while if in a database of 100 names a 1% misidentification occurs, which equals one person, the system would appear to be working perfectly, but if the database has a million persons, a 1% rate or error would result in ten thousand mistakes. A random effects model, however, assumes that the present database is a sample from a still bigger database and as a result inference easily extends to the bigger database. The method operates on the premise that a database of authentication results from the existing system already exists, and that with various hierarchical random effects models, and Bayesian inference techniques, posterior predictive distributions can generate a prediction of error rates in the biometric system (Mitra et al., 2007).

This method allows one to predict outcomes of biometric authentication systems when their use is expanded to larger groups of peoples, or different groups of people. The method also better predicts the probability of a false alarm in an evaluation (Mitra et al., 2007). The model was tested on three different face authentication systems, a filter-based system, a Gaussian Mixture Model based system and a frequency domain representation of facial asymmetry system. The overall significance of the study was that the proposed methodology "provides an alternative means of performance evaluation to those based on empirical observations studies by providing model-based prediction" (Mitra et al., 2007, p. 30). Though Mitra et al. (2007) tested the model on face recognition systems, they also proposed that the method can be adapted for use with any biometric systems, if one has the match scores from a database.

The switch to multimodal biometrics.

Argyropoulos et al. (2010) noted that of various biometric strategies to combat identity theft the integration of two or more biometric traits, so-called multimodal biometric systems, have increasingly become a subject of interest. Multimodal systems have been found to be able to overcome many of the limitations imposed on identity fraud prevention by unimodal biometric systems. Unimodal biometric systems have a number of problems in capturing clean data, including problems related to noisy data, intra-class variations, restricted degrees of freedom, non-universality, spoof attacks and unacceptable error rates. By using a multimodal system, capturing different types of biometrics to contribute to an overall identity profile, many of these problems can be circumvented. This is because integrating data from more than one source allows for verification of identity according to the more stringent performance requirements of a multimodal system (Nanavati, Thieme, & Nanavati, 2002). Another strength of multimodal systems is that they can make use of any number of independent biometrics, allowing them to

work around limitations of unimodal devices based on data limits in the population (Mane & Jadhav, 2009).

Some of the problems that unimodal systems face, which multimodal systems can work around, include the fact that 3% of people have illegible fingerprints, colds and sickness often alter voices, invalidating voice identification systems, and face recognition is notoriously difficult to make due to incalculable changes in ambient light and the pose of the subject at the moment of data recording. All of these problems can be overcome when a multimodal system feeds various biometric data into a system, matching data in ways that cancel out errors resulting from unimodal problems (Mane & Jadhav, 2009).

Another strength of multimodal as opposed to unimodal biometrics is that it is much more difficult to forge multimodal biometric characteristics compared to unimodal. Thus, while a person may be able to forge facial recognition or fingerprint modes, for the person to be aware of the different kinds of data required by a multimodal system is much more challenging. As such, then, "multimodal biometrics is….much more resistant to fraudulent technologies" (Argyropoulos et al., 2010, p. 164). This is called biometric system robustness, which makes the system resistant to fraudulent attack. Attacks against biometric systems can originate from any number of places. A fraudulent identity attempt to breach security can occur at the sensor, in the database, or at the sensor, or at the communication channel between the sensor and the database and matcher, as well as at the point of output.

Wechsler (2010) adopted an information processing perspective on biometrics, by way of evaluation. He found that face recognition in uncontrolled or even moderately controlled environments is still problematic. To improve biometrics, multimodal biometrics have been introduced, as have new attempts to fuse the data derived from multimodal biometric acquisition. At present, however, Wechsler found that much of the data fusion taking place is ad-hoc and imprecise, and that research must advance an integrated, principled and unified methodology for biometric inference using randomness and complexity concepts.

Wechsler (2010) proposed "a novel all encompassing methodology for robust biometric inference and prediction built around randomness and complexity concepts" (p. 510), designed to improve the fusion of multimodal biometric data. Thus, it is more accurate to state that multimodal biometrics supported by identity management will result in much improved biometrics.

Sulovska and Adamek (2010) reviewed the various kinds of biometric systems, to determine which is most useful at the present moment. Again, they argued that biometrics generally represent a step forward from the use of cards or passwords because of "the permanent holding of biometric character by its living carrier and minimal opportunity for stealing it by a trespasser" (p. 1463). The testing focused on the fact that most biometrics can still be overcome because few have as yet installed liveness protection, that is, fake forms of biometrics can be presented.

With regard to face recognition, Sulovska and Adamek (2010) argued that those systems making use of artificial neural network are most advanced because they can work around obstructions like sunglasses. They acknowledged iris recognition for its potential, especially since the iris itself does not change over time and varies greatly from individual to individual. Color iris recognition has also been recently developed, providing for more system improvement. Fingerprinting is still favored because it is least disturbing the public and fingerprint readers are widespread (Sulovska & Adamek, 2010). But fingerprint systems can also be rather easily

attacked, necessitating the creation of new algorithms. The fact that fingerprinting has advanced to include liveness detection is also important.

Sulovska and Adamek (2010) evaluated the effectiveness of fingerprint biometrics by calculating the ROC ratio, again the ratio between false accept match rate and false rejection rate. Partial results found that if one fakes a fingerprint with an adhesive from a fusion gun it can still defraud the system. Overall, at this point in ongoing research, Sulovska and Adamek concluded "it appears useful to use a multi-biometric system in order to prevent incommodity during identification" (p. 1464). Though at present the only common multimodal systems involve the use of both a biometric and a PIN number, any number of possibilities exists. Overall, in reviewing face, iris and fingerprint recognition, Sulovska and Adamek appeared to have greatest hope for iris recognition but at present concede that each unimodal method continues to have vulnerabilities that perhaps could only be reduced by developing a multimodal biometric security system.

Volner and Bores (2009) conducted an evaluation of some multiple biometric combination systems. Facial or iris recognition, which seems to be more easily accepted by the public than fingerprinting, was studied. A personal encounter with multimodal systems at Heathrow Airport was described, primarily using face and iris recognition. Volner and Bores found that, in terms of public usability, or ease of use, the current crop of multimodal biometric systems are limited by the fact that the subjects must work to position themselves correctly to get a good reading. While systems with automatic height adjustment and auto focus have been developed, and with better results, current practice must deal with the positional limitation. Volner and Bores also concluded "the technical community should focus more on helping people manage identity management and their own biometrics" (p. 59).

Mane and Jadhav (2009) concurred that many unimodal biometric systems suffer from enrollment programs to the non-universality of biometric traits, making them susceptible to biometric spoofing or insufficient accuracy caused by noisy data. Mane and Jadhav argued that most unimodal systems, which account for most biometrics systems currently in commercial operation, are also vulnerable to variations and spoofing, and are known to have high false acceptance rates (FAR) and false rejection rates (FRR). If, however, multiple biometrics were taken, many of the limitations of unimodal systems could be overcome. Multimodal approaches, for example, solve the problem of universality because multiple traits ensure wider population coverage. Spoofing is also blocked as it becomes more and more difficult for an impostor to spoof multiple biometric traits. As a result, they prefer multimodal biometric authentication systems, which combine several different biometric measures to arrive at a decision about a person's true identity (Mane & Jadhav, 2009). Among the multimodal biometric models developed thus far, in the multi algorithm approach a single biometric sample taken from the sensor is processed using a number of different algorithms (Jain et al., 2004). They cited the example of the 2002 Face Recognition Vendor test that found that performance accuracy increased when 2D face recognition biometric results were combined with results from different commercial recognition systems (Mane & Jadhav, 2009). Another type of multi-model biometric is multi sample algorithms that make use of multiple samples of the same biometric. Each sample is processed and the results fused to obtain an overall recognition result. The processing in multimodal approaches, examples of which were reviewed, can be more complicated. This approach has been found to overcome the liability of poor samples, but it does require higher expense for sensors and increased cooperation from the user providing multiple samples. At present, though expense appeared to act as a hindrance as did actual technology

development, Mane and Jadhav argued that the improved integration of multiple sensors, scalability improvements and quality measures to quicken decision making will all eventually make multimodal biometrics the norm. Nonetheless, because "performance gain is pronounced when uncorrelated traits are used in a multimodal system" (p. 92), it is Mane and Jadhav's view that multimodal biometrics are superior to unimodal biometrics and will in due time replace the current biometrics methods being used today.

Finally, Gorodnichy (2009) contrasted the development of biometric systems—primarily from unimodal systems operating in controlled environments to multimodal systems that can work in a variety of uncontrolled environments--to the fact that evaluation of biometric systems has remained unchanged. Most biometric evaluation is still measured based on reports of false match and non-match rates as cited above and the tradeoffs in the ROC curve that results. While this method has become, previous studies have documented, the norm, Gorodnichy argued that this kind of evaluation may "no longer be sufficient and appropriate for investigating the performance of state-of-the-art systems" (p. 1).

As a result, a gap has opened up between theory and practice in the expert evaluation of biometric systems. Gorodnichy (2009) argued that what he termed multi-order performance analysis can be applied to obtain all-inclusive descriptions and evaluations of biometric performance. To describe this analysis method, Gorodnichy reviewed the means by which biometric evaluation evolved, starting with two research areas, image processing, derived from computer science, and pattern recognition, derived from statistical machine learning theory.

Gorodnichy (2009) also argued that there are two stages of biometric deployment, the passage data stage, when new biometric images are presented to the computerized system, and the enrollment stage, when the new images are matched with those already stored in the system. The operational recognition tasks of most biometric systems involve verification, identification, screening, classification and similarity quantification, all of which in one way of another measure the biometric input for universality, uniqueness, permanence, performance, collectability and acceptability. A trade-off between performance and acceptability is the norm in basic evaluation. For these operations to be undertaken operational conditions are imposed entailing dichotomies of overt versus covert, cooperative and not, structured or not, local versus centralized and the relative impact of false matches versus those of false non-matches.

That said, biometric recognition may still fail, and evaluation is all about determining why a biometric reading failed (Introna & Nissenbaum, 2006). The evaluation will have to look into the capture of the image, how they were enhanced, what kind of feature detection was undertaken, the computation of the template, the computation of match scores, which recognition decision is used (the most common being a binary comparison to a fixed threshold.

An error in any of these steps of a biometric system may result in the poor recognition of the biometric (Introna & Nissenbaum, 2006). An evaluation of all of these steps asks questions like is the image quality good, are the used features informative, is the iris extraction good and what is the confidence level of the extraction? Evaluation would then make a suggestion of returning to image processing and pattern recognition basics to determine why the system failed (Introna & Nissenbaum, 2006). With the addition of biometric systems using multimodal techniques to extract data in surveillance-like environments the recognition results increasingly need to be integrated or fused over time with results obtained from other biometric readings (Jenkins & Burton, 2008).

As biometrics and video surveillance merge, with systems having titles such as Biometrics on the Move or Biometrics on the Go being developed, this overlap will place greater

demands on evaluation (Jenkins & Burton, 2008). Thus, Gorodnichy (2009) presented a case study where to evaluate the effectiveness of various new multimodal face recognition systems, one needs to know what distinguishes this kind of biometrics from other biometrics. In so-called stand-off biometrics, when the person does not directly interact with the sensor, and sometimes not even know that biometric data is being extracted, an evaluation can only be made based on multimodal biometric measurements taken from more than one sensor using some data fusion technique (Gorodnichy, 2009).

Instead of relying on the concept of security involving the opening or closing of a door, correctly or incorrectly, resulting in false match and false non-match rates and the resulting tradeoff ROC curve, "a new evaluation framework needs to be developed that allows one to obtain the all-inclusive description of the performance of a biometric system based on its place in biometric taxonomy and all data measured during the run of the system" (Gorodnichy, 2009, p. 4). That is, the false accept and false reject instances must be expanded into cumulative measures with various proposed metrics or curves developed, as indicated by Gorodnichy (2009). An all-inclusive evaluation would involve determining the suitability of the modality, whether or not costs have impacted the FM or FNM, determine all factors affecting performance and evaluating the performance of various market solutions. As a result, a multi-order analysis is undertaken, which can be used to fine-tune the system and improve confident recognition results.

Overall, Gorodnichy (2009) concluded, "the recognition performance needs to be understood and all performance changes that are due to a change of a system or system parameters and not only the match/non-match errors have to be analyzed" (p. 18). In today's world, any organization relying on biometric technology to safely conduct business not only required constant monitoring of the biometric system performance, but a regular all-inclusive system performance evaluation. Just as biometrics itself is evolving from unimodal to multimodal, so too expert evaluation is quickly being transformed from a statistical evaluation based on a single ratio to a systemic discipline taking into consideration all of the complicating factors that can compromise a multimodal biometric system.

CONCLUSION

This review examined the usefulness of biometrics as means of combating identity fraud in numerous different fields (Gregory, 2008; McCollum, 2005; Mercuri, 2006; Opperman, 2009; Smith & Lias, 2005; Swartz, 2009; Winterdyk & Thompson, 2008). The review briefly surveyed the current state of identity fraud prevention efforts, based mostly on improving organizational or public attention to the issue, including the introduction of red flags. Most identification continues to be authenticated through passwords or ID cards of some sort. As a result, many researchers have proposed biometrics as a solution to the problem of rising identity fraud. In studies of biometrics, however, a number of researchers have noted barriers to its implementation in its current technological state. While biometrics is theoretically preferred to other identify fraud security efforts, the research at present suggests that cost and public receptiveness issues may still be holding back its development (Dey & Samanta, 2010; Douhou & Magnus, 2009; Kriemer, 2010; Leong & Yerzak, 2004; Palaniappan, 2008; Sullivan, 2009).

Nonetheless, a good many studies have assessed the degree to which biometrics generally improves security, and how particular types of biometrics, from fingerprint and iris identification, to keystroke and even voice biometrics, also improve security (Kim & Bzullak, 2008; Li & Wechsler, 2009; Nikam & Agarwal, 2009). When it comes to the evaluation of

biometric systems, some challenges remain (Coventry, 2005; Dekking & Hensbergen, 2009; Gorodnichy, 2009; Jain & Pankanti, 2001; Mane & Jadhav, 2009; Schuckers, 2003; Sulovska & Adamek, 2010; Volner & Bores, 2009; Wechsler, 2010). That said, a number of studies tested specific biometric systems and found the security they offered superior to password or ID card-based security.

In comparing currently implemented biometric systems, researchers appear to theoretically favor iris identification but acknowledge that the lack of technology and expertise needed to make that form of biometrics more acceptable to the public continue to hold it back. As such, using the technology acceptance model, it appears that fingerprint biometric elements installed in all manner of security machines at present were preferred. It is also true that many researchers, having compared various methods of biometrics, have concluded that comparing different types of biometrics will not in itself lead to improvement, but that biometrics in general must move from a unimodal to a multimodal platform of development.

To this end, Mane and Jadhav (2009) claimed that any single biometrics will still, no matter how much it is improved, not deliver an optimal outcome, and still have security problems. The answer to this problem is to develop multimodal biometrics that takes in two or more types of biometric data, making it almost impossible to defraud (Mane & Jadhav, 2009). In addition to this, most evaluations of biometrics are currently based on the ROC curve, based on the ratio of the system trade-offs having to due to false test occurrences. Some researchers have also argued that as biometrics moves to a multimodal platform more complicated evaluation rubrics will also need to be developed (Kleist, 2007).

Overall, then, the review found that experts prefer fingerprint biometrics to all other unimodal biometrics, but that they are also shifting their support to multimodal systems with evaluation undertaken by more complicated means than the current ROC measure (Heckle, Patrick, & Ozok, 2007). Thus, this literature review found in biometrics a quickly developing field that is continuing to evolve at a rapid pace, with expert evaluations of individual biometric systems being quickly overtaken by a broader theoretical shift to multimodal biometrics.

REFERENCES:

Argyropoulos, S., Tzovaras, D., Ioannadis, D., Damousis, Y., Strintzis, M. G., Braun, M., & Boverie, S. (2010). Biometric template protection in multimodal authentication systems based on error correcting codes. *Journal of Computer Security, 18*, 161–185. doi:10.3233/JCS-2010-0369

Ashbourne, J. (2004). *Practical biometrics: From aspiration to implementation*. London, England: Springer-Verlag.

Coventry, L. (2005). Usable biometrics. Retrieved from http://hornbeam.cs.ucl.ac.uk/hcs /teaching/GA10/lec3extra/ch10-1coventry.pdf

Dekking, M., & Hensbergen, A. (2009). A problem with the assessment of an Iris identification system. *Society for Industrial and Applied Mathematics Review, 51,* 417–422.

Dozier, K. (2010). *Yemen official: Arrested woman didn't mail bombs.* Retrieved from http://www.aolnews.com/world/article/yemen-official-arrested-woman-didnt-mail-bombs/19696677

Goldwasser, J., & Anderson, T. M. (2007). Passwords + pictures = security? *Your Money Kiplinger's,* June, 79–80.

Gorodnichy, D. O. (2009). Evolution and evaluation of biometric systems. *Proceedings of the IEEE Symposium on Computation Intelligence for Security and Defense Applications.* Retrieved from http://videorecognition.com/doc/publications/09-cisda-evol-eval-P.pdf

Gregory, A. (2008). Conserving customer value: Improving data security measures in business. *Journal of Database Marketing and Customer Strategy Management, 15,* 233–238. doi:10.1057/dbm.2008.20

Hong, J. H., Yun, E. K., & Cho, S. B. (2005). A review of performance evaluation for biometrics systems. *International Journal of Image and Graphics, 5,* 501–536.

Hoonakker, P., Bornoe, N., & Carayon, P. (2009). Password authentication from a human factors perspective: Results of a survey among end-users. *IEEE Symposium on Security and Privacy, 53,* 459–463. Retrieved from www.ieee-security.org/TC/SP2011/PAPERS /2011/paper003.pdf

Jain, A., & Pankanti, S. (2001). Biometrics systems: Anatomy of performance. *Journal of IEICE, E00-A,* 1–11.

Jain, A. K., & Pankanti, S. (2008). Beyond fingerprinting. *Scientific American, 299,* 1–6.

Jarillo, G., Pedrycz, W., & Reformat, M. (2008). Aggregation of classifiers based on image transformations in biometric face recognition. *Machine Vision and Applications, 19,* 125–140. doi:10.1007/s00138-007-0088-9

Kleist, V. (2007). Building technologically based online trust: can the biometric industry. *Information Systems Management*, *24,* 319–329.

Mane, V. M., & Jadhav, D. V. (2009). Review of multimodal biometrics: Applications, challenges and research areas. *International Journal of Biometrics and Bioinformatics, 3,* Retrieved from http://www.cscjournals.org/csc/manuscript/Journals/IJBB/volume3 /Issue5/IJBB-33.pdf

Mansfield, A. J., & Wayman, J. L. (2002). *Best practices in testing and reporting.* Retrieved from http://homepage.ntlworld.com/avanti/bestpractice.pdf

McCollum, T. (2005). Flaws found in identity protection. *Internal Auditor, August,* 20–21.

Mercuri, R. T. (2006). Scoping identity theft. *Commmunications of the ACM, 49,* 17–23.

Mitra, S., Savvides, M. & Brockwell, A. (2007). Statistical performance evaluation of biometric authentication systems using random effects models. *Pattern Analysis and Machine Intelligence*, *29*(4), 517-530. http://eksl.isi.edu/files/papers/sinjini_2007_1172280675.pdf, p. 1-34.

Nanavati, S., Thieme, M., & Nanavati, R. (2002). *Biometrics: Identity verification in a networked world*. New York, NY: John Wiley & Sons.

Pinheiro, R. (2004). Preventing identity theft using trusted authenticators. *Journal of Economic Crime Management*, *2*, 1–16.

Podio, F., & Dunn, J. (n.d.). *Biometric authentication technology: From the movies to your desktop*. Retrieved from http://www.itl.nist.gov/div893/biometrics /Biometricsfromthemovies.pdf

Schuckers, M.E. (2003). Using the beta-binomial distribution to assess performance of a biometric identification device. *International Journal of Image and Graphics, 3,* 523-529.

Stana, R. M. (2002). *Identity fraud: Prevalence and links to alien illegal activities* (GAO-02-830T). Retrieved from http://www.gao.gov/new.items/d02830t.pdf

Sulovska, K. & Adamek, M. (2010). Research on Biometrical systems: an overview. Annals of DAAAM for 2010 & Proceedings of the 21st International DAAAM Symposium, 21, 1463-1464.

Swartz, N. (2009). Will red flags detour ID theft? *Information Management, February,* 38-41.
Tappert, C.C., Cha, S.H., Villani, M. & Zack, R.S. (2010). A keystroke biometric system for long-text input. *International Journal of Information Security and Privacy, 4,* 32-60. DOI:10.4018/jisp.2010010103
Volner, R. & Bores, P. (2009). Biometric techniques in identity management systems. *Electronics and Electrical Engineering, 7,* 55-59.
Wayman, J., Jain, A., Maltoni, D., & Maio, D. (2005). *Biometric systems, technology, design and performance evaluation*. New York, NY: Springer.
Wechsler, H. (2010). Intelligent biometric information management. *Intelligent Information Management, 2,* 499-511.
Wilcox, N. A., Jr., & Regan, T. (2002). *Identity fraud: Providing a solution*. Retrieved from http://www.lexisnexis.com/risksolutions/IdentityFraudWhitePaper.pdf