

Using fingerprint image quality to improve the identification performance of the U.S. Visitor and Immigrant Status Indicator Technology Program

Lawrence M. Wein*[†] and Manas Baveja[‡]

*Graduate School of Business and [†]Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305

Edited by Burton H. Singer, Princeton University, Princeton, NJ, and approved March 8, 2005 (received for review October 8, 2004)

Motivated by the difficulty of biometric systems to correctly match fingerprints with poor image quality, we formulate and solve a game-theoretic formulation of the identification problem in two settings: U.S. visa applicants are checked against a list of visa holders to detect visa fraud, and visitors entering the U.S. are checked against a watchlist of criminals and suspected terrorists. For three types of biometric strategies, we solve the game in which the U.S. Government chooses the strategy's optimal parameter values to maximize the detection probability subject to a constraint on the mean biometric processing time per legal visitor, and then the terrorist chooses the image quality to minimize the detection probability. At current inspector staffing levels at ports of entry, our model predicts that a quality-dependent two-finger strategy achieves a detection probability of 0.733, compared to 0.526 under the quality-independent two-finger strategy that is currently implemented at the U.S. border. Increasing the staffing level of inspectors offers only minor increases in the detection probability for these two strategies. Using more than two fingers to match visitors with poor image quality allows a detection probability of 0.949 under current staffing levels, but may require major changes to the current U.S. biometric program. The detection probabilities during visa application are ≈ 11 – 22% smaller than at ports of entry for all three strategies, but the same qualitative conclusions hold.

biometrics | homeland security | policy | queues | game theory

The September 11, 2001 attacks may have been prevented if some of the 18 terrorists were apprehended as they entered the U.S. (1). To rectify this situation, the multibillion dollar U.S. Visitor and Immigrant Status Indicator Technology (US-VISIT) Program (2) takes two index fingerprint images from each visa applicant and matches these prints against those of several hundred million visa holders to detect whether the new applicant already has a visa under a different identity (3). Visitors (i.e., visa-holders, as well as citizens of 27 visa-exempt countries who are visiting for >90 days or are traveling on work or student visas) to U.S. ports of entry have two new fingerprints taken, which are compared to the original prints to verify that they are who they claim to be, and are used to identify whether visitors are on a watchlist that contains known criminals and suspected terrorists. The US-VISIT Program also uses a facial image for verification (i.e., one-to-one matching); although face recognition technology has improved over the last several years, it is not a viable tool for identification (i.e., one-to-many matching) from a pool of millions (4). Here, we develop a mathematical model to compare various biometric identification strategies at visa application and at ports of entry.

Identification, which is much more difficult than verification, is performed by software systems that compute a similarity score between any two fingerprints. Upon entry to the U.S. (a similar procedure is used during visa application), a visitor is fingerprinted during primary inspection and is assigned a pair of similarity scores, one for each index finger, against each person in the watchlist. If either the left or right score exceeds a given threshold, or the sum of the left and right scores exceeds a second

threshold, then the corresponding person on the watchlist is added to the visitor's candidate list. If the candidate list for a visitor is nonempty, then this visitor is further investigated during a secondary inspection. We call a visitor illegal if he is on the watchlist and legal otherwise. If a visitor is illegal, then he will have a set of fingerprints (taken at an earlier point in time), called a mate, in the watchlist. The detection probability (called the true accept rate in ref. 5) is the probability that an illegal visitor's mate is placed on the candidate list (i.e., the similarity scores between his new set of fingerprints and his mate exceeds one of the thresholds). The false positive probability (called the false accept rate in ref. 5) is the probability that a legal visitor's candidate list is nonempty. For two index fingers and for the fingerprint databases tested by the National Institute of Standards and Technology (NIST), the identification system currently used in the US-VISIT Program had a detection probability of 0.959, independent of the watchlist size, and the false positive probability increased with the size of the watchlist and was 3.1×10^{-3} when the watchlist had 6 million people (5), which is comparable in size to the actual watchlist during entry. To avoid undue congestion at ports of entry under current staffing levels, it is necessary to maintain the false positive probability at approximately this level.

The identification system used in the US-VISIT Program also computes the image quality of a fingerprint, which reflects the inherent quality of the print and operational factors such as humidity, dirt, and finger pressure. NIST has determined that the identification performance of the biometric system is highly dependent on the fingerprint image quality, with better performance resulting from good-quality images (5, 6). The present study stems from the belief that terrorist organizations can exploit the image quality-dependent performance of the biometric identification system by choosing from their large pool of potential U.S.-bound terrorists, those that have either inherently poor image quality (e.g., worn out fingers) or deliberately reduced image quality (e.g., surgery, chemicals, sandpaper). We formulate and solve a Stackelberg game (7) in which the U.S. Government chooses the parameters for a biometric identification strategy to maximize the detection probability subject to a constraint on the mean total (i.e., primary plus secondary) biometric processing time per legal visitor, and then the terrorist chooses the image quality level to minimize his detection probability. We use this optimization problem to assess three types of strategies, including the strategy currently used in the US-VISIT Program. One of these strategies requires additional primary biometric processing, which is why our constraint is in terms of

This paper was submitted directly (Track II) to the PNAS office.

Freely available online through the PNAS open access option.

Abbreviations: US-VISIT, U.S. Visitor and Immigrant Status Indicator Technology; NIST, National Institute of Standards and Technology.

[†]To whom correspondence should be addressed. E-mail: lwein@stanford.edu.

© 2005 by The National Academy of Sciences of the USA

Table 1. A description of the three biometric strategies

Strategy	Parameters	Condition for placement on candidate list	Optimal solution	Detection probability
Current	t_1, t_2	$\bigcup_{j=1}^2 (s_j > t_1) \cup (\sum_{j=1}^2 s_j > t_2)$	1317, 1818	0.526
Two-finger	$t_{11}, \dots, t_{18}, t_{21}, \dots, t_{28}$	$\bigcup_{j=1}^2 (s_j > t_{1j}) \cup (\sum_{j=1}^2 s_j > t_{2i})$	4741, 3690, 3010, 2853, 3257, 1977, 1743, 1120, 5263, 4936, 4989, 4783, 3831, 2647, 2303, 1120	0.733
Multifinger	$t_1, \dots, t_8, n_1, \dots, n_8$	$\bigcup_{j=1}^{n_i} (s_j > t_i)$	1975, 1804, 1678, 1512, 1820, 1573, 1901, 1378, 2, 2, 2, 2, 3, 4, 5, 10	0.949

The first two strategies use two fingers from each visitor, whereas the multifinger strategy uses n_i fingers if the visitor has quality $i = 1, \dots, 8$. The values of (n_1, \dots, n_8) can be based on quality information obtained at visa enrollment. The similarity scores between the visitor and each watchlist person are (s_1, s_2) for the first two strategies, and (s_1, \dots, s_{n_i}) for the multifinger strategy if the visitor has quality i . The third column describes the condition on the similarity scores for placing a watchlist person onto a visitor's candidate list, where the "U" denotes "logical or." The last two columns give the values of the parameters that solve Eqs. 1–3, and the corresponding optimal detection probability in Eq. 1, assuming a watchlist size of 6 million. The "Current" strategy is used in the US-VISIT Program.

the mean total biometric processing time per legal visitor rather than the false positive probability.

The Model

In our model (see *Supporting Text*, Figs. 3–7, and Tables 2–6, which are published as supporting information on the PNAS web site), each person has an associated image quality that is an integer value between 1 (highest quality) and 8 (lowest quality) (5). This random quantity is assumed to be independent and identically distributed across people, whether they are visitors or on the watchlist. This assumption ignores the fact that right and left fingers of the same person may vary in quality, operational noise may cause a specific finger to have a different image quality when retested, and people's fingers may wear out over time, causing the image quality of an illegal visitor to be worse than that of his mate on the watchlist. Although our model and analysis can be generalized so that each image, rather than each person, can possess a different quality, more detailed unpublished data from NIST would be required to parameterize the model. However, the analysis in Fig. 3, which uses data from ref. 5 and the belief that the level of operational noise can be kept low via good training and processes, suggests that this assumption holds true in the great majority of cases and can be made without affecting our qualitative conclusions.

The two fundamental building blocks of our model are the intraperson similarity scores, which quantify the match between an illegal visitor's fingerprints with the earlier pair of prints in the watchlist, and the interperson similarity scores, which compare a visitor's fingerprints with a different person's prints from the watchlist. Because each person is assumed to have a given quality level, the intraperson similarity scores are described by a family of eight probability distributions, one for each quality level. Following ref. 8, we assume that an interperson similarity score depends on the qualities of the two matched fingerprints only via the worse of the two qualities, which allows us to consider only eight interperson similarity score distributions. Guided by quality-aggregated data from NIST (9), we use gamma distributions for intraperson scores and log-normal distributions for interperson scores (see *Supporting Text* for details).

We consider the three biometric strategies described in Table 1, which are referred to as the current strategy, the two-finger strategy, and the multifinger strategy; the two-finger strategy allows quality-dependent thresholds and the multifinger strategy allows the thresholds and the number of fingers tested to vary according to quality. It takes ≈ 5 s per finger to take an image of a visitor's fingerprint (3, 10) and ≈ 2.5 s per finger to perform the software matching against the entire watchlist (10). We assume that the watchlist matching is performed in parallel with a

visitor's interview at the port of entry, so that matching does not increase a visitor's processing time (10); however, the physical fingerprinting process needs to be managed by the inspector to guarantee low operational noise and to detect fraud (e.g., artificial fingerprints). In addition, we assume that 20 min are required on average to perform a secondary inspection for each false positive, regardless of the size of the candidate list. Let d_i be the detection probability if the illegal visitor has image quality i , and let f denote the overall false positive probability; these are computed in *Supporting Text* for each of the three strategies. If we let m_1 be the mean time to image a single fingerprint, m_2 be the mean secondary inspection time of positive matches, and $p(i)$ be the fraction of visitors that have quality i for $i = 1, \dots, 8$, then the mean total biometric processing time per legal visitor is $m_1 \sum_{i=1}^8 p(i)n_i + fm_2$ for the multifinger strategy, and $2m_1 + fm_2$ for the other two strategies. Substituting the values $m_1 = 5$ s, $m_2 = 20$ min, and $f = 3.1 \times 10^{-3}$, we find that the base-case value of the mean total biometric processing time per legal visitor for the two two-finger strategies is $2m_1 + 3.1 \times 10^{-3}m_2 = 10 + 3.72 = 13.72$ s at the port of entry. The optimization problem for the multifinger strategy is

$$\text{maximize minimize}_{i=1, \dots, 8} d_i \quad [1]$$

$$\text{subject to } m_1 \sum_{i=1}^8 p(i)n_i + fm_2 \leq 13.72, \quad [2]$$

and Eq. 2 is replaced by

$$f \leq 3.1 \times 10^{-3} \quad [3]$$

for the other two strategies. The maximization in Eq. 1 is carried out over the parameters in the second column of Table 1, assuming a watchlist of 6 million people at the port of entry.

Results

The use of quality-dependent thresholds on the similarity scores increases the detection probability from the current level of 0.526 to 0.733 (Table 1). This strategy achieves the same detection probability for all quality levels by using a low threshold ($t_{28} = 1120$) for the worst quality (the single-finger threshold t_{18} is redundant in this optimal strategy). The multifinger strategy achieves a detection probability of 0.949 by using 10 fingers for the lowest quality, and 3–5 fingers for the next three lowest quality levels.

The detection probability can be improved by increasing the mean biometric processing time per legal visitor, i.e., by increas-

two-finger strategy, from 0.733 to 0.775, while maintaining existing congestion levels at the ports of entry. As of May 2004, US-VISIT Program plans do not call for additional staff or facilities at land ports of entry (2). Testing 10 fingers rather than two fingers for poor-quality visitors can increase the detection probability at the U.S. border to 0.949 under current staffing levels. Moreover, during visa application, the two-finger strategy can achieve a detection probability of only 0.569 if the false positive probability is set at 3.1×10^{-3} , whereas a 10-finger strategy can achieve a detection probability of 0.840. However, the US-VISIT Program only takes images of two fingers during visa application (5), despite previous warnings that a two-finger system was inadequate for identification with large watchlists (6). Although switching from two to 10 fingers at this point in time, even for only poor-quality images, may be expensive and disruptive (6), this multifinger approach appears to be a more cost-effective exception-management alternative for poor-quality images than other biometrics (e.g., iris, retina, hand geometry; see ref. 3) or human interrogation. Slower, more accurate matching techniques for poor-quality images should also be assessed and compared to the multifinger approach. The extent to which these options are pursued should be assessed in light of the detection probability required to deter terrorists from attempting a border crossing at an official port of entry, which itself depends on the terrorists' perceived likelihood of successfully entering the U.S. between the ports of entry, e.g., along the U.S. borders with Mexico and Canada. The detection probability between the ports of entry on the U.S.–Mexico border has been estimated at 0.25 (12), although it appears that, at this point in time, Al Qaeda prefers to enter the U.S. at ports of entry (1).

The quality-dependent biometric analysis performed here has other potential applications. First, a terrorist could intentionally deface his fingerprints between the time his watchlist fingerprints were imaged and the time he enters the US-VISIT Program. If the fingerprints are only partially altered, then the low thresholds associated with poor-quality images in our two-finger strategy (Table 1) should increase the likelihood of detection slightly, and the multifinger strategy would significantly hinder the terrorists' success with this approach. This topic deserves further investigation. Another possible application is to assess the degradation in quality of faxed fingerprint images, which appears to have been at the crux of a mistaken terror arrest in Spain (13).

Although our qualitative conclusions are likely to be robust, the quality-dependent intra- and interperson similarity score distributions at the core of our model were indirectly estimated from quality-dependent detection probability vs. false positive probability curves. The use of unpublished NIST data on the

intra- and interperson similarity scores broken down by image quality, and the actual watchlist, which likely generates worse performance than publicly available databases (5), perhaps coupled with a more refined model that allows quality to be associated with an image rather than with a person, would be required to sharpen our policy recommendations and to derive operationally reliable parameter values. Among other parameter values in our model, only the mean secondary processing time (m_2 in Eq. 2) has not been reported in the literature. This parameter only affects the relative performance of the multifinger strategy, because it is the only strategy that trades off primary and secondary inspection. We may be underestimating the performance of the multifinger strategy because the four non-thumb fingers can perhaps be imaged simultaneously (14), which would reduce the mean primary processing time for a 10-finger print from 50 s to 20 s.

In conclusion, there appears to be a serious but reparable vulnerability (detection probability is highly dependent on image quality) in the biometric identification system of the US-VISIT Program, which is the last, and perhaps main, line of defense for keeping terrorists off of U.S. soil. Our analysis provides the government with a means of assessing the worst-case threat, and there is a silver lining in the equilibrium solution, namely, that the resulting detection probability will be equal for all image quality levels, rendering the system more robust than the existing system. This, in itself, is a strong reason for switching to the proposed policy. The introduction of quality-dependent thresholds requires only minor software modifications and can increase the detection probability by ≈ 0.2 , and should be implemented as soon as possible. The use of more than two fingers for low-quality images can increase the detection probability to ≈ 0.95 ; in other words, the worst-case performance under the proposed multifinger strategy is approximately the same as the existing strategy's performance under the naive assumption that terrorists do not behave strategically at all. There is no excuse for a multibillion dollar program to settle for performance below the level of the proposed multifinger strategy, particularly given the potentially grave consequences of a false negative. Our policy recommendations hinge on the assumption that terrorist organizations will attempt to defeat the biometric system by employing terrorists with poor-quality fingerprints. In light of the meticulous planning that has gone into terrorist attacks over the last decade (1), we believe this assumption is not only prudent, but realistic.

We thank Michael Garris for sharing data from refs. 5 and 9. This research was supported by the Center for Social Innovation, Graduate School of Business, Stanford University, and by a fellowship from the Center for International Security and Cooperation, Stanford University.

1. National Commission on Terrorist Attacks (2004) *The 9/11 Commission Report* (Norton, New York).
2. U.S. General Accounting Office (2004) *First Phase of Visitor and Immigration Status Program operating, but improvements needed* (Government Accountability Office, Washington, DC), report GAO-04-586.
3. U.S. General Accounting Office (2002) *Using Biometrics for Border Security* (Government Accountability Office, Washington, DC), report GAO-03-174.
4. Phillips, P. J., Grother, P., Michaels, R. J., Blackburn, D. M., Tabassi, E. & Bone, J. M. (2003) *Face Recognition Vendor Test 2002* (National Institute of Standards and Technology, Gaithersburg, MD), National Institute of Standards and Technology International Report 6965.
5. Wilson, C. L., Garris, M. D. & Watson, C. I. (2004) *Matching Performance for the US-VISIT IDENT System Using Flat Fingerprints* (National Institute of Standards and Technology, Gaithersburg, MD), National Institute of Standards and Technology International Report 7110.
6. The Attorney General, Secretary of State & the National Institute of Standards and Technology (2003) *Report to the Congress: Use of Technology Standards and Interoperable Databases with Machine-Readable, Tamper-Resistant Travel Documents* (Office of the Attorney General, U.S. Department of State, and National Institute of Standard and Technology, Gaithersburg, MD).
7. Gibbons, R. (1992) *Game Theory for Applied Economists* (Princeton Univ. Press, Princeton).
8. Tabassi, E., Wilson, C. L. & Watson, C. I. (2004) *Fingerprint Image Quality* (National Institute of Standards and Technology, Gaithersburg, MD), National Institute of Standards and Technology International Report 7151.
9. Wilson, C. L., Watson, C. I., Garris, M. D. & Hicklin, A. (2003) *Studies of Fingerprint Matching Using the NIST Verification Test Bed (VTB)* (National Institute of Standards and Technology, Gaithersburg, MD), National Institute of Standards and Technology International Report 7020.
10. Edmunds, T., Sholl, P., Yao, Y., Gansemer, J., Cantwell, E., Prosnitz, D., Rosenberg, P. & Norton, G. (2004) *Simulation Analysis of Inspections of International Travelers at Los Angeles International Airport for US-VISIT* (Lawrence Livermore National Laboratory, Livermore, CA).
11. Halfin, S. & Whitt, W. (1981) *Oper. Res.* **29**, 567–588.
12. Bartlett, D. L. & Steele, J. B. (September 20, 2004) *Time*, Vol. 164, No. 12, p. 51.
13. Kershaw, S. (June 5, 2004) *N.Y. Times*, Section A, p. 1.
14. Hicklin, R. A. & Reedy, C. L. (2002) *Implications of the IDENT/LAFIS Image Quality Study for Visa Fingerprint Processing* (Mitretek Systems, Falls Church, VA), www.mitretek.org/publications/biometrics/NIST-IQS.pdf.