FAQ: Biometric Encryption Paper

Q. What is Biometric Encryption (BE)?

- Biometric Encryption is a technology that allows you to use your biometric to encrypt a PIN, a password, or an alphanumeric string, for numerous applications – to gain access to computers, to enter buildings, and to privately and securely prove identity, etc.
- The PINs can be 100s of bits in length; the length doesn't matter because you don't need to remember it!

Q. What problems does BE address?

- Loss of individual control over one's personal (i.e., biometric) data.
- Data matching, secondary purposes, function creep, surveillance, profiling.
- Loss, theft, misuse and abuse of personal data (e.g. identity theft, fraud).
- Compliance with privacy and data protection laws.

Q. What are the benefits of BE?

- No need to retain biometric data = NO central database.
- Improved individual control: unlinkable and revocable identifiers, encryption made easy.
- No need to memorize long passwords.
- Improved security: data; authentication; communications.
- Greater confidence, trust, and use of biometrics by the public.

Q. How ready is BE technology for deployment?

- The technology has matured considerably since the mid-1990s. Interest has accelerated since 2002. BE is now close to the prototype stage.
- Similar technologies are currently being developed around the world. These are noted in the paper.
- Especially promising is work being done by Philips Research, which have systems operational and ready for deployment.

Q. Where/how could BE be used / deployed?

- BE can be deployed anywhere that "traditional" biometrics can be used, i.e., for authentication and access control, and to encrypt data.

- The paper provides three case scenarios: one small-scale use (biometric authentication); the second a medium-scale one (medical database access); the third, a large-scale (national ID card).

- BE is especially well-suited for use with smart cards and mobile devices. BE can be used to create biometric tickets, such as personalized boarding passes.

Q. Don't BE-enabled systems suffer from the same weaknesses as other biometric systems?

- All biometric systems share some generic weaknesses, such as the need for a trusted enrolment process, spoofing concerns, a trusted and secure infrastructure (e.g. readers), as well as performance reliability issues.

- But BE moves the privacy and security yardstick forward considerably on many fronts. These benefits are documented in our paper.

- Additional security can always be added to BE, such as requiring a smart card and/or password/PIN.

Q. Why did you write this paper? What are you trying to accomplish?

- We are now at a critical inflection point for a widespread take-up of biometric-enabled information systems across society.

- We have long advocated for building privacy early and directly into information technologies and systems, wherever possible.

- Privacy and security are not opposites and should not need to be "traded-off."

- We want to swing the spotlight on a true "win-win" privacy-enhancing technology (PET) – to show that privacy-enhanced biometrics are possible, and should be considered and supported.

- We want to encourage further research and support for this exciting technology, and to help BE develop to its full potential before it's too late.

Q. How does BE address the privacy concerns associated with national identity cards and travel documents?

- BE makes possible strong and secure local authentication to the identity card, without recourse to a central database of biometric data. If a check against a central database is necessary, BE makes it far more secure.

- Case Study #3 in the paper describes a scenario involving use of BE-enabled ID cards at a border crossing.

- As Philips Research have shown, it is possible to apply facial recognition BE to travel documents, and to engage 3-way verification in a privacy-enhanced manner.

Technical Questions

Q: How much entropy is there in a fingerprint?

A: The issue about fingerprint entropy has been controversial for a long time. There is no definite answer to it. The most comprehensive paper published to date is by Pankanti, S., Prabhakar, S., and Jain, A, titled, *On the individuality of fingerprints*. IEEE Transactions on PAMI 24 (2002), p.p. 1010-1025.

There are two basic approaches to the problem: a) theoretical modeling; and b) empirical estimation.

a) Theoretical Modeling

An example of very simplistic theoretical modeling is the following: Let us consider a typical 500 dpi resolution fingerprint image of 300x300 pixel size. It contains 30 minutiae which are uniformly and randomly distributed and are statistically independent of each other. Each minutiae is characterized by 4 parameters: x-location, y-location, angle, and type (either ridge ending or bifurcation). To accommodate for intra-class variations, we have to provide some windows of tolerance. If it is 20 pixel for x and y locations, there will be 15x15 windows across the 300x300 image. In other words, each x and y can be quantized to 4 bits. We can also quantize the angle to 4 bits, i.e. with the window for angle 360/16 = 22.5 degrees. One bit for the type. In total, we get 30x(4+4+4+1) = 390 bits. This estimation is, probably, too optimistic, as the randomness and independence assumptions do not hold very well.

A more realistic model was presented in the foregoing Pankanti et al paper. For a fingerprint containing 36 minutiae, the authors estimated a probability that two fingerprints will falsely match on all 36 minutiae as 5.5×10^{-59}. Therefore, the entropy is $-\log_2(5.5 \times 10^{-59}) = 193$ bits.

It should be noted that most models do not take into account non-minutiae information (i.e., the ridge pattern) of fingerprints. However, contrary to what has been stated for decades, a non-minutiae fingerprint verification algorithm outperformed all minutiae algorithms in two international competitions in a row, FVC2002 and FVC2004. In other words, the non-minutiae information should also have a significant entropy, although it is very difficult to estimate it.

b) Empirical estimations are usually based on John Daugman's paper on iris recognition

J. Daugman, *High confidence visual recognition of persons by a test of statistical independence*, IEEE Trans. Pattern Anal. Mach. Intell. 15 (11) (1993), p.p. 1148–1161.

An inter-class (i.e. impostor's) distribution of verification score for a large data sample is computed. The histogram of the distribution is approximated with a normalized binomial distribution. The parameter N of the distribution is called a number of degrees of freedom. For iris, J. Daugman obtained N = 249 bits.

The problem with this approach is that it is very algorithm and database dependent – the results are affected by intra-class variations and the algorithm tolerance. The binomial distribution does not approximate well the "tail," i.e., where all false acceptances occur.

An example of this empirical estimate for fingerprints can be found in a paper published by P. Tu and R. Hartley titled, *Statistical Significance as an Aid to System Performance Evaluation*, Lecture Notes in Computer Science, Springer, Volume 1843/2000, p.p. 366-378.

For a 100-bit minutiae code (representing only a 100x100 pixel part of the image), the authors found the entropy to be 82 bits, i.e. a moderate loss of 18 bits. We observed the same tendency: if the extracted features are properly randomized and binarized, the entropy loss is usually less than 20%.

In general, the empirical results should be interpreted with caution.

In summary, we think that a realistic number for the fingerprint entropy lies between 200 and 300 bits. This is still an area of ongoing research.

Q: What kind of encryption keys can be bound to fingerprints?

A: We target a key of 128 bits or so to be bound to the fingerprints. Experiments in the past showed that this is feasible. Even if the fingerprint entropy is not sufficient, we can try a multimodal approach, i.e., using two or more fingerprints or iris/face + fingerprint. As we show in the paper, high distortion variability of fingerprints makes this task more challenging than for the iris or even the face, (note that Philips Research has the Biometric Encryption based on face ready for deployment).

Q. Don't all biometrics, including BE, suffer from the "glass slipper" effect? i.e., a given individual's biometric will be capable of revealing a given key forever, while other people's biometrics won't. This being the case, how will BE offer any advantage in preventing future mining of databases for biometric matches?

A. Let us consider a not-so-distant future scenario. When the use of biometrics grows, an ordinary person will be enrolled in various biometrically controlled databases, such as travel documents, driver licenses, health care, access control, banking, shopping, etc. The current, (i.e. conventional, non-BE), biometric systems can use the same biometric template for all of them. The template becomes the ultimate unique identifier of that person. This is where biometric data mining comes into effect: the different databases, even if some of them are anonymous, may be linked together to create comprehensive personal profiles for all the users. To do this, no fresh biometric sample is even required. The linking of the databases can be done offline using template-to-template matching, in a very efficient one-to-many mode. The privacy implications explode at this point.

Contrast that to BE: it would be much more difficult, if not impossible, to engage in the linkage of biometric databases. BE does not allow a template-to-template matching – the tool commonly used in conventional biometrics. In each BE database, a user has different keys bound to his biometric. Those templates cannot be matched against each other. You need a real biometric sample to do so. Moreover, this matching is relatively slow and, therefore, highly inefficient in one-to-many mode. For example, running a single image against 10,000,000 records in just one BE database could take 0.1 sec x 10,000,000 = 1,000,000 sec = 11.5 days.

It is basically correct to state that if an individual's real biometric image was somehow obtained, then this "glass slipper" could be used to search various databases for all the different PINs or keys that "fit" and, accordingly, construct a personal transaction profile of the individual concerned, using data mining techniques. But you would first have to obtain a "satisfactory" real image of the correct biometric and/or multiple biometrics used to encrypt the PIN or key. All of the PINs or keys in the databases can and should be unique (the privacy in numbers argument) – as such, if an individual's actual biometric could somehow be accessed, only an ad hoc data mining search could be made, accessing only one entry, (which would represent an individual privacy breach, not a breach of the entire database).

However, with BE, the actual biometric (or template derived from that biometric) is never stored – a record of it doesn't exist. Without the actual biometric, data mining techniques would be useless because there would be no common template to use as one's search parameter. As mentioned, all the biometrically encrypted PINs or keys in the databases would be unique. Furthermore, access to the individual's biometric and associated transaction data would be far more difficult if a biometrically encrypted challenge/response method is employed.

In contrast, current biometric methods use a common (the same) biometric template for an individual's transactions and, accordingly, can be used as the search parameter to construct personal profiles, without access to the real biometric. This presents both a privacy and security issue because not only could profiles be constructed on an ad hoc basis, but each template in a database can be used to construct profiles of multiple individuals without access to their real biometric. We thus believe that this alone makes biometric encryption far superior to standard current biometric methods.