GRESHAM COLLEGE

22ND OCTOBER 2019

BIOMETRICS: HOW UNIQUE ARE YOU?

PROFESSOR RICHARD HARVEY FBCS

I'm writing this lecture at Mumbai airport while on my way home to Amsterdam from Hyderabad. So, far, and I have not yet boarded my aircraft, my passport and boarding pass have been checked twenty-five times! At each presentation a security guard, member of the Indian armed forces, or an airline employee has gazed intently at my passport and my boarding pass, looking up each time, to check that I match the photograph in my passport. At one point a member of the Indian Army checked my passport twice to see if I could be trusted to mount an escalator[1]. Of course, one might moan that India is in love with senseless bureaucracy or, as we shall see in this lecture, possibly India's long history with biometrics makes it particularly prone to such excessive zeal when checking documents. However, the UK, the situation is equally challenging. Consider the case of someone who falls on hard times who is ejected from their house for not paying the rent. Their landlord, in a fury burns their possessions. They have no cash and need to get some from the bank but have no bank card? How can they obtain one? Well a replacement bank card needs an address to which it can be sent. If the address changes then the bank may well require sight of a photographic identity document. So, to get a bank card we need a driving licence or a passport. How can we get a driving licence? We need a passport. How do we get a passport? Simple, obtain your parents' birth certificates, two signed photographs, pay £85, wait six weeks, have an interview and then you have a passport. The you can get a diving licence (another £34) and now you can get a bank account. Seems extreme? Far from it, this is a real situation. If it had not been for the kindness of friends, this person would have starved to death while the UK Home Office shuffled the passport paperwork.

Biometrics, measurements from the body, hold out the hope that those twenty-five document checks will be reduced to none. They allow for the possibility that interactions with the State will be smooth and error- and fraud-free.

The problem is that only one or two biometrics work to the level that people might expect them to work. Leaving aside DNA, Iris scans and possibly fingerprints, most biometric systems are toys which are probably more secure than a pin code but little more than that.

How easy it to guess a PIN code? There are four digits in a PIN code, so an arbitrary guess is a 1 in 10,000 chance. Most machines allow three guesses before locking your card so roughly 3 in 10,000 chance. However, pin codes are compromised all the time. Partly because people tend to use codes which are simple to guess but mostly because it is quite feasible to persuade people to reveal their code. One way is to video someone at the keyboard. Even if this does not reveal precisely the code, any partial knowledge is enough to bring down the 3 in 10,000 chance to 3 in 1000 (one digit revealed); 3 in 100 (2 revealed); 3 in 10 (3 revealed). Not bad odds! Another way is to build a fake cash point. One clones the card and records the PIN code and bingo!

When I got my first bank account, one used to draw money by going into the bank and writing a cheque. The bank clerk would receive the cheque, they would check the signature against a "bank card" which looked like a credit card but just had a copy of one's signature on the back. If they matched and the clerk felt it was you, you were given money. What was less well known was that, if one forgot one's chequebook or bank card, then one could write a "counter cheque" which, provided the bank clerk knew you (or someone knew you), was just as

[1] In my case I could not – I was too early to mount this particular escalator and was sent back to an outer circle of hell to wait.

effective. In effect one relied on the bank-clerk as a face-detector. But how accurate are people when identifying other people?

Such things have been studied extensively by psychologists and it's not good news. In one experiment [1], a number of shoppers were equipped with fake photo-credit cards. Roughly half the cards were accepted as genuine despite the shop assistants being incentivised to spot them. In another experiment [2] a tourist with a map asks for directions from a Local on a university campus. Three people carrying a door barge between the tourist and the local – during the fractions of a second that the door separates them, the tourist is swapped for one of the people carrying the door. Over half the people did not notice the switch. It appeared that some people did not notice even when the ethnicity of sex of the tourist changed! Thus, not only are ordinary humans not good at recognising faces, they are not good at recognising change in faces.

On the face of it, this is good news for computer scientists working in biometrics; the benchmark to beat, provided by humans, is ludicrously low. Unfortunately, while it is useful to know that humans are terrible at recognition of humans, especially if one is accused of murder, the standards demanded of computers are much higher. Computer software can be replicated without cost, so the consequences of inaccurate recognition are much more serious than those of say a slightly wonky bank clerk.

Notwithstanding the bizarre faith placed in human recognition by police officers and the courts, there is a long-running countermovement which asserts that crime investigation would be improved by accurate measurements of humans or biometrics. The initiation of this movement is generally credited to Alphonese Bertillon [3]. The son of a statistician, Bertillon, was determined to bring order to the world of forensic identification – his anthropomorphic system involved the making of systematic measurements of the head and body. These were recorded alongside two photographs of the person (side-one and frontal) – a configuration now known as the "mug shot". Bertillon is also credited with the invention of the fingerprint although it would be fairer to say that he developed a systematic approach to finger printing which came to prominence in a 1902 murder trial in France (the first use of fingerprints in the court room)[2].

Before turning to biometrics, we ought to review the process of user authentication. Early systems were based on something that users knew, but others did not, such as a password, a username or a PIN or all three[3]. These systems are still commonplace and many of us are familiar with the hassle of generating codes that are long enough to be difficult to guess, but memorable. An alternative method, which also has as a very long a history, is the *token* (often a card, a dongle or a key). Although no memory is required (although one has to remember one's keys of course), token-based systems suffer from the disadvantage of easy copying and of the implicit, and erroneous, assumption that whoever has the key is authorised to use it. Of course, both systems can be combined, as with a cash-card (a card plus a PIN).

Biometrics represent a third approach – here we use something unique about a user. A good biometric should have the following properties:

- Universality. Every human being on the planet should have this feature. Gait recognition, for example, assumes that everyone can walk which is clearly not true. The vast majority of people do have fingerprints though[4].
- Uniqueness. Every person should have a unique version of this biometric. Think identical twins to see if you can work out why some biometrics are not likely to work.

---

[2] Although Bertillion appears to have been a rather eccentric individual this did not stop Sir Arthur Conan Doyle referring to him in several Sherlock Holmes novels with some admiration.

[3] NatWest's corporate credit card system demands that you remember all three. It also asks for only certain digits, out of order, to be produced within a certain time limit. On your way home from this lecture, see if you can write down the 13th, 5th,10th and 3rd characters from the passphrase "Xn126vbGhj132" in under 10 seconds. "Which?" magazine consistently ranks NatWest poorly in its response to fraud.

[4] The condition of having no fingerprints is known as *Adermatoglyphia*. It has not been much studied because it is so rare but the first paper I could find on this was called "The immigration delay disease: Adermatoglyphia–inherited absence of epidermal ridges" by Burger, Bettina *et al.* in theJournal of the American Academy of Dermatology, Volume 64, Issue 5, 974 – 980.

- Permanency. A biometric that ages is a nuisance. Your height for example changes quite noticeably during your lifetime.
- Collectability. What technology is required to collect this biometric? Can it be measured in a timely fashion? With modest amounts of computation?
- Acceptability. This is a most tricky aspect. There are large cultural differences in what might make an acceptable biometric and society appears to go through spasms of unacceptableness – not all of them logical.

A nice table listing some common biometrics is given in [4] and reproduced below:
Table 1: Common biometrics listed by property (adapted from [4])

	Universal?	*Unique*	*Permanent?*	*Collectible?*	*Acceptable?*
DNA	Yes	Yes	Yes	Tricky	Poor
Iris	Yes	Yes	Yes	Yes	Marginal
Fingerprint	Yes	Yes	Yes	Yes	Fair
Face	Yes	Yes?	No	Yes	Good
Voice	No	Yes	No	Yes	Good
Gait	No	No	No	Yes	Good
Keystrokes	No	No?	No	Yes	Good

It's important to realise that some of the columns in Table 1 are, not only a matter of opinion, but also very dependent on context. For example, the collection of fingerprints might well be regarded as intrusive were it to be mandatory at your bank, but it is now routine at border-crossings so now counts of merely one of a series of indignities associated with international travel.

Although biometric systems vary greatly, a system model for biometric systems might look something like Figure 1.

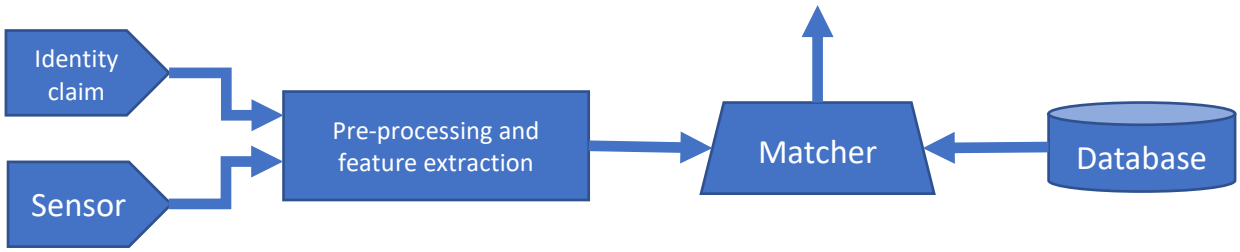

Figure 1: *Biometric schematic modified from Ratha, Nalini & Connell, Jonathan & Bolle, Ruud. (2001).* Enhancing Security and Privacy in Biometrics-Based Authentication Systems. *IBM Systems Journal. 40. 614-634.*

Figure 1 shows a hybrid system in which an individual provides a biometric (via the lower channel) but also a claim for identity (via the upper channel)[5]. The first part, the feature extraction, is different for each biometric. The matching process, the second part, can need quite sensitive implementation as, firstly the database often needs to be kept very secure, and secondly large-scale matching needs to happen quickly.

At this point we should mention a rather delicate issue which is that many biometrics are not very good. There are very considerable commercial rewards for successful biometrics so there is considerable commercial pressure to pretend that the performance of a system is acceptable when it is not. Furthermore, as we have already seen, it is not difficult to produce a biometric that is better than an average human operator recognising faces. This problem of snake-oil biometrics was realised quite early on in the history of biometrics, but it was not until the

[5] The Privium system as Schiphol airport is a hybrid system. At enrolment a used has their iris scanned and encrypted onto a card. In use the user places their card in a reader, then has their iris scanned. If the card code matches their iris then they enter the airport bypassing passport control.

1990s that there was a concerted effort to design standardised tasks. In summary there is always a trade-off between security (being sure we accept only genuine people) and usability (not rejecting too many genuine people). If just one figure is quoted, then it is often the case that the other figure is not so quotable! One solution is to fix the system to have Equal Error Rate (EER) in which case the two probabilities are equal, and we can summarise performance with just one number. More generally scientists would ask that we place costs on the effects of errors. In my experience users are quite extraordinarily reluctant to do this[6].

However, when it comes to attacks on biometric systems, this is still an emerging field[7],[8] and computer scientists may not be best placed to think of all the naughty ways that criminals or malicious people might behave. Indeed, various authors have pointed out that every point in Figure 1 is a potential security breach waiting to happen.

Notwithstanding these problems, widescale adoption of biometrics is on its way. Probably the largest scale system is known colloquially in India as the *Aadhaar* system which is administered by the Unique Identification Authority of India (UIDAI). The system has roughly 1.2Bn enrolled and is used to administer the Indian Social Security system which is explains its popularity. The UIDAI dashboard current claims a total of around 40M authentications per day. Unfortunately, there have been repeated rumours that UIDAI is not secure. An example of one of these reports can be found in, say [5]. It is reported that the breach arises because "photo on each record page used the file name as that worker's Aadhaar number, a confidential 12-digit number assigned to each Indian citizen as part of the country's national identity and biometric database," but later in the article it states "Aadhaar numbers aren't strictly secret, but are treated similarly to Social Security numbers." So, are they secret or not? To answer that question is complex as one has to think through all the consequences of knowledge of an item. And one needs to consider whether widespread knowledge of that item is particularly more damaging that restricted knowledge.

This leads us to the final point about biometrics – this is one of the areas where society needs to be involved and be involved in a way that understands the nuances and subtleties of biometric identifiers. In Europe the recent regulation of data via the General Data Protection Regulations 2018 specifically mentions biometric data as a "sensitive category of personal data" so requires special care. This means that holders of such data are, firstly, discouraged from holding it and, secondly, are required to conduct a "privacy impact assessment" before holding such data[9]. Not all countries have the equivalent and in India, one of the prime movers in biometrics, the law is laxer and the Aadhaar data breach, if it was a breach, appears to have happened without much punishment. The USA functions via state laws of course, California appears to be in the vanguard with the California Consumer Privacy Act (CCPA) and Washington State has recently introduced a state privacy bill. In short, legal frameworks are being developed for handling biometric data but certainly there is a big gap between what the public understands and what the public wants.

In the meantime, a little common sense will have to go a long way.

---

[6] At a recent Board meeting of an organisation that shall remain nameless I was told that the target for a system was zero errors. When I questioned the tolerable level for errors, I was told it was zero. The fact that this was impossible did not seem to bother my interlocutors.

[7] A recent review paper [3] defined five new probabilities, or rates, associated with presentation (or PAD) attacks. As an example, consider BPCER (Bona Fide Presentation Classification Error Rate) which is proportion of bona fide presentations which are incorrectly identified as PAD attacks.

[8] I was writing this section just after listening to a fascinating mini lecture by Professor Rainer Böhme from the University of Innsbruck. He noted that modern computer systems consistently fail to understand that users have conflicting interests. He notes that were we to move away from the utopian vision of the "perfect user" then computer systems would be a lot more secure.

[9] As somebody once remarked – one can hardly change one's fingerprints when some contractor has a data breach.

1. *When Seeing should not be Believing: Photographs, Credit Cards and Fraud.*, R Kemp, N Towell and G Pike, Applied Cognitive Psychology. Jun1997, Vol. 11 Issue 3, p211-222
2. *Failure to detect changes to people during a real-world interaction*, Simons, D. J., & Levin, D. T. (1998), Psychonomic Bulletin and Review, 5, 644–649.
3. *Presentation Attack Detection for Iris Recognition: An Assessment of the State-of-the-Art*, Adam Czajka and Kevin W. Bowyer. 2018. ACM Comput. Surv. 51, 4, Article 86 (July 2018), 35 pages. DOI: https://doi.org/10.1145/3232849
4. *An introduction to biometric recognition,* Anil Jain, Arun Ross and Salil Prabhakar, IEEE Trans on Circuits and Systems for Video Technology, Vol 14, No 1, Jan 2004, pp 4 – 20.
5. https://techcrunch.com/2019/01/31/aadhaar-data-leak/