Topic 22: User identification

George Markham
Reg 100130020
george.markham@uea.ac.uk

December 5, 2018

Abstract

Due to the increasing amount of sensitive data people entrust to online services it is important that this data is kept secure and only accessible by authorised parties. Usernames & passwords consistently fail as verification and identification methods therefore alternatives must be considered. This report discusses both biometric and behavioural biometric methods of verification and identification. The methods by which researchers determine their performance is also investigated. This report finds that both are good candidates for replacing usernames & passwords and, while biometric identification is likely better in most cases than behavioural biometric identification, for verification tasks behavioural biometrics are both less obtrusive and can work as well as their biometric counterparts.

Contents

1 Introduction

User identification has become increasingly important with the advent of the internet. Businesses, governments, and the general public are sharing personal, sometimes very sensitive information with trusted parties. The key issue being that word *'trusted'*. How does one prove one's identity to another user or system across the globe?

There are seemingly constantly news stories relating to data theft, last year's Equifax (credit reporting agency) hack could have potentially leaked nearly half the population of the United State's extremely sensitive data (Ng & MUSIL 2017). The amount and the sensitivity of the data held on most of the general population of the world can be incredibly damaging if leaked, it can potentially lead to serious fraud that can have very damaging consequences for people. This data theft can be hindered by using secure ways of identifying and authenticating users. Some of those methods will be discussed in section 2. If implemented correctly they can potentially mitigate the risk of data theft. Some methods of more secure identification are already being implemented, for example companies such as Google, Facebook and PayPal allow the use of Two Factor authentication, where a user augments the security of their username and password by entering a separate code sent to their device (usually a mobile phone). The code adds extra security as a potential attacker is unable to access a user's account and data unless they also have access to the user's phone. These codes are also resistant to brute force attacks as they are only valid for a short amount of time.

For many years we have relied on passwords to authenticate and identify ourselves but passwords are consistently proven to be insecure and unreliable forms of identification. There are, however, alternatives to the username and password method. Smartphones are including biometric identification methods such as using fingerprints sensors and facial analysis. Other methods for identification can also be considered, for example typing habits, mouse use and speaker recognition can all potentially be used to uniquely identify a person.

1.1 Background and Key Issues

The main issue with non-biometric based identification methods is that they are easily forgeable. Take the example of a fingerprint, even amongst identical twins you are still able to be identified with no significant decrease in accuracy (Han et al. 2004). Compare that with the assumption that a password can be known to many different people (either legitimately or illegitimately) then it is clear that biometric identification methods could be superior to traditional username and password methods.

Biometric methods encompass many different ways to identify a user. Sections 2.2 and 2.3 will cover the various methods of using both biometric (fingerprints, iris scanning etc...) and behavioural biometric techniques (keystroke dynamics, mouse dynamics, speaker recognition etc...) to identify a user. The usefulness of these techniques varies with the use case, for example it is illogical to attempt to use typing habits to identify a user on a telephone call, equally it is not always feasible to use voice identification when authenticating a user for a website.

Alternative methods will also be discussed including the possibility of identifying a user with real-world documents such as a passport, drivers license or ID card. These systems are already secured against fraud so they may prove to be a great candidate for use with computer systems.

In addition to the methods mentioned above there are also public-key encryption algorithms already developed and in use that, although developed for sending encrypted messages, could also be considered forms of identification.

1.2 Aim and Objectives

This study aims to provide evidence that biometric, behavioural biometric and alternative forms of user identification are superior to a traditional username and password based method. To do this the various methods of identification outlined in section 2 will be evaluated using evidence gathered from previous studies to determine their usefulness compared with each other and username and password identification. Biometric and behavioural biometric forms of identification will be compared to highlight the merits of each.

1.3 Study Plan

As part of this report the various methods of identification will be compared and contrasted in order to outline their respective strengths and weaknesses. These comparisons will be based on data from past work external to this study. Each technique will be compared to other, similar, techniques and their various merits and weaknesses will be discussed in 3 and 2. Examples of both biometric and behavioural biometric methods will be discussed in order to give a broad overview of the current state of both fields.

2 Methods

2.1 Usernames & Passwords

Usernames and Password authentication is extremely common. It is the primary way most large websites and apps like Facebook, Google, Instagram etc... authenticate their users, however it has been proven to be rather insecure. The main factor making passwords insecure is that the user must select their own and remember it. This leads users to use the same password across multiple systems, select short, and therefore easy to crack, passwords or to select passwords that are personal to them. If a user uses the same password across a number of different systems then if one system is compromised, or stores the password in an insecure manner then the user can potentially have multiple compromised accounts with relatively little effort by malicious parties. If a user uses a short password, e.g 4 alphanumreic characters, then it would take a system capable of calculating 1 million hashes per second about 15 seconds to crack that password (Kessler 1996). If the user selects a password with just 4 lowercase letters then it could take approximately 0.5 seconds to crack. This is trivial for a malicious party and the kind of computing power needed to achieve this is easily obtainable.

2.2 Biometrics

The use of biometric forms of identification eliminates many of the issues surrounding traditional username and password identification. For example the issues surrounding users selecting weak passwords is eliminated as one does not control one's biometric features and therefore cannot weaken the security of that system. Biometric forms of identification are also more unique and almost impossible to brute force. However there are still security issues with biometric forms of identification. One can attempt to trick the sensor for example by creating a fake finger or face (Ambalakat 2005). It is possible to extract finger print patterns from an image taken by a normal digital camera (Ogane & Echizen 2017), however it is also possible to protect against such an attack, researchers created *"BiometricJammer"* to "effectively prevent the illegal acquisition of fingerprints by surreptitious photography" (Ogane & Echizen 2017). The jamming pattern proposed, while effective, must be worn by a person. This means that users must take active action in order to protect themselves. Other methods of biometric identification could be used instead to offer more protection to users. Iris recognition provides similar distinctiveness and performance to fingerprint recognition however it may appear more invasive so is unlikely to be accepted as a widespread form of identification. (Ambalakat 2005).

2.3 Behavioural Biometrics

Behavioural biometrics pertains to biometric features that are not physical attributes rather patterns in behaviour including speech and typing habits. Just as biometrics can be used to uniquely identify a person through physical attributes so can behavioural biometrics through that person's behaviour. Often this can be done unobtrusively, for example when accessing a website if one could be identified by the keystrokes leading up to that website visit then it eliminates the need for that person to actively prove their identity. By utilising potentially more secure methods of authentication both a user's experience and security can be enhanced. There are a number of different behavioural biometrics to be considered and many of these will be discussed in section 3.

2.4 Process for Identifying or Verifying a User

The standard process for identifying a user based on behavioural biometrics is much the same as the one used for normal biometrics. It relies on feature extraction, a database of templates, and a way of matching features to a particular user. For enrolment of a user the biometric is read by a sensor, checked for it's quality and then passed to a feature extractor (Jain et al. 2004). These features can then be stored in a database for identification and verification purposes (figure 1). Verification and identification follow much the same process, however remove the quality checker and instead move straight into feature extraction. For verification the user's template is found in the database and then a matching process attempts to match that template with the new extracted features. If the match is successful then it can be assumed that the user is who they say they are. To identify a user out of a number of possible users the biometric system will select a number of templates and try to verify a user againstt them until it finds a match, if it doesn't find a match then it is assumed the user is not known to the system (Jain et al. 2004).

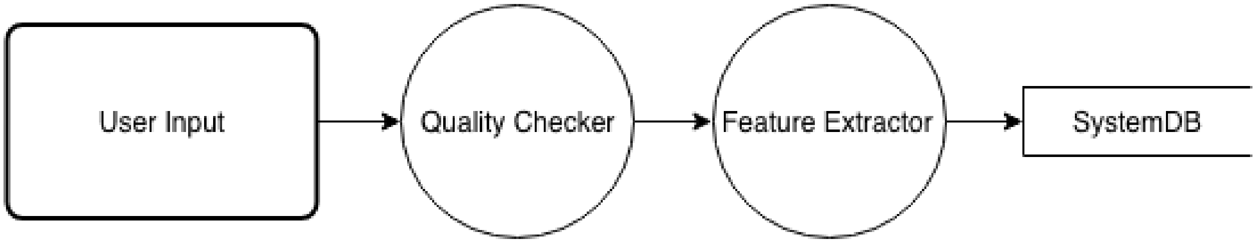

Figure 1: Flowchart Showing User Enrolment In A Biometric System (Jain et al. 2004)

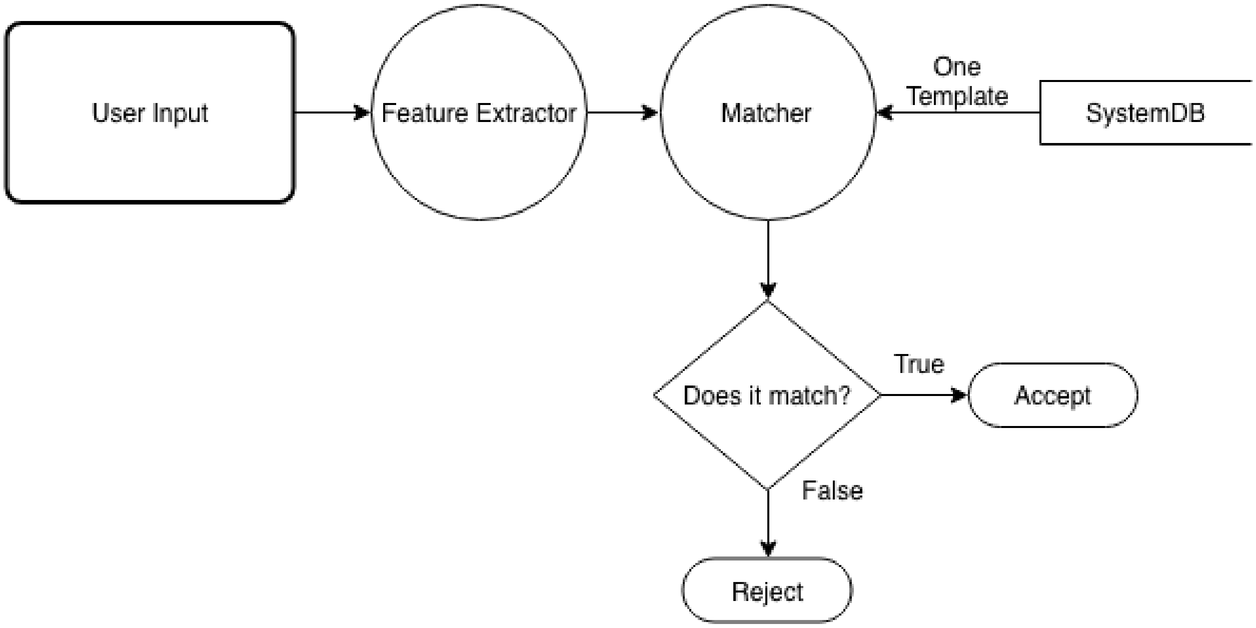

Figure 2: Flowchart Showing User Verification In A Biometric System (Jain et al. 2004)

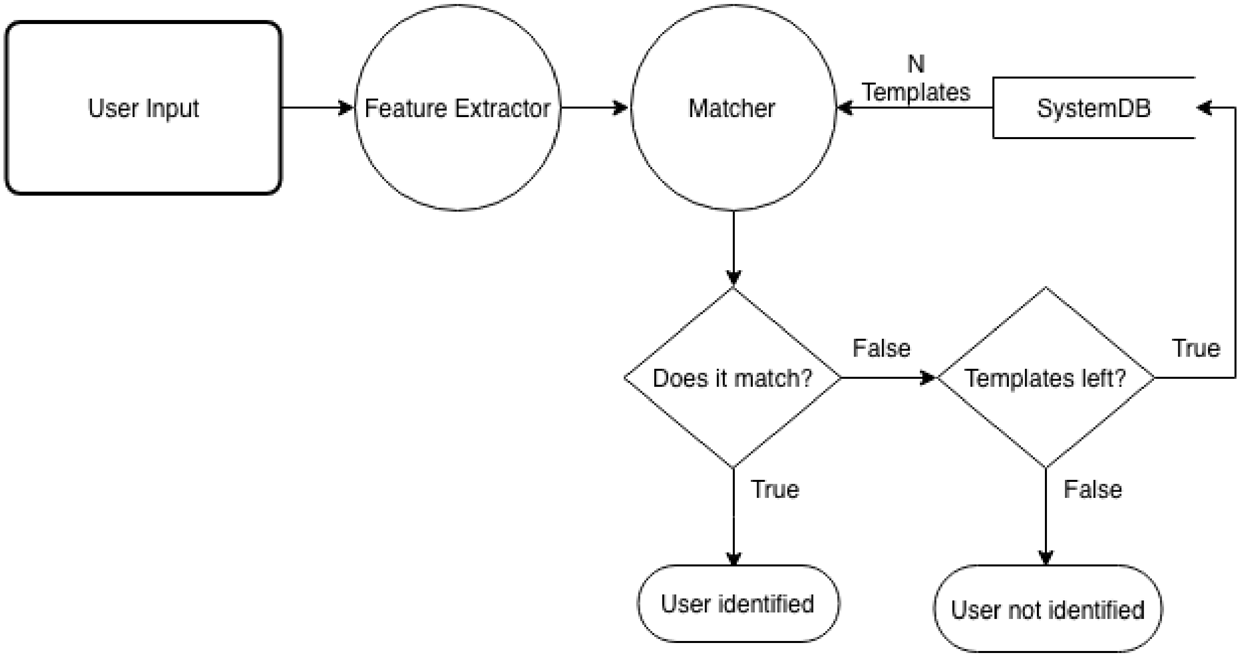

Figure 3: Flowchart Showing User Identification In A Biometric System (Jain et al. 2004)

2.5 Alternative Methods

Other methods of identification include the use of government verified documents such as drivers licences and passports. Trusona Inc. have developed a system that allows users to securely identify themselves with a physical ID card (Eisen 2017). Trusona has created authentication systems for the FBI (Abagnale 2017) thus highlighting how secure this technology can be. However this technology does have a significant barrier to entry, a user needs to have a physical form of identification in order to identify themselves on the system. In the cases of undocumented immigrants or those who do not have documents such as passports, drivers licences or other forms of physical identification it would not be possible for them to use this system.

2.6 Specific Methods

2.6.1 Fingerprints

Fingerprint identification is widely used in smart phones such as Apple's iPhone and Samsung's S series of devices. Given it's wide use it's important that it is doesn't identify two different people as the same person, measured by *false match rate* (FMR) (Delac & Grgic 2004). The FMR of fingerprint identification has been observed as 0.2% (Delac & Grgic 2004). This is very low however the experiment only appears to include adults between the age of 20-39, excluding many younger and older people who also need access to fingerprint identification technology. It is not sufficient to include a subset of the population when discussing identification technology. Another issue with fingerprint technology is when user's do not have fingerprints. Fingerprint loss can occur through some medical treatments including chemotherapy. When it comes to smartphone authentication and identification technology it may be prudent to explore alternative biometric methods that are more universal with lower barriers to entry like face identification (Jain et al. 2003).

2.6.2 Face Identification

As mentioned above face identification is now being introduced into Apple's new iPhone models (X, XS and XS Max). Traditionally face identification has faced technological issues, specifically a lack of accuracy in "environments with cluttered backgrounds and varied lighting conditions" (Jain et al. 2003). Apple's *"FaceID"* technology avoids this issue by projecting 30,000 infared dots onto the user's face to produce a sequence of depth maps and infared images (Apple 2017). To further enhance security and avoid spoofing the sequences are randomised (Apple 2017).

2.6.3 Handwriting Recognition

As a method of user identification handwriting recognition is arguably more accessible and useful today as opposed to before the advent of widespread, accurate touchscreen devices. There are two types of handwriting recognition: *on-line* or *dynamic* and *off-line* or *static* (Tappert et al. 1990). Off-line handwriting recognition is processed after a person has written something. For example a person may have written a letter on paper, this could then be analysed by a handwriting recognition system and the person's identity could be determined. Because identification does not happen in real-time off-line handwriting recognition is more suited to multi-factor authentication systems, where a user may provide one form of identification initially and then be required to back that initial identity challenge with an example of their handwriting. More useful today, perhaps, is on-line handwriting recognition, where a user may be identified as they are writing using a transducer (Tappert et al. 1990). This is of note given the large number of touchscreen devices such as phones, tablets, laptops and even some desktop computers that can be used as inputs to handwriting recognition systems.

2.6.4 Signature Recognition

Quite closely related to handwriting recognition (2.6.3) is signature recognition. Similar to handwriting recognition there are two methods of performing the identification, *on-line* and *off-line*. Signatures are already widely used in many aspects of people's lives, this reduces the barrier to entry as most people will already be used to signing documents and have standard way of doing it. It is also worth noting that people can find all aspects of unfamiliar technology quite frightening and by using methods already familiar to people, like signature recognition, it may make user's feel more at ease with the system.

There are issues with signature recognition, people are liable to change their signatures over a period of time, may vary substantially and are liable to forgeries (Jain et al. 2004). These are substantial issues when attempting to securely identify a person. Variations make can make systems liable to false matches as two similar signatures with in-built variance may appear to be the same. Signature recognition is also liable to spoof-attacks (Jain et al. 2004). Traditional methods of signature forgery will still apply to signature recognition systems and, given that signature forgery has been used to defraud people for many years, it can be assumed that this practice will render signature recognition system insecure for identification over a network.

2.6.5 Keystroke Dynamics

Given that one of the main inputs to most computer systems is a keyboard (either a software keyboard on a touchscreen or a hardware keyboard) it would be useful if one could be identified through it. Using features extracted from a user's typing, including time-features such as down-down times, down-up times, up-down times and the key code of the key typed by the user (Araujo et al. 2005). These features in the system proposed in Araujo et al. (2005) gave false rejection rates of 1.45%, false acceptance rates of 1.89% for impostor users and false acceptance rates of 3.66% for impostors that had observed legitimate users typing habit. This was proposed as an augmentation to traditional username and password systems to make them more secure without requiring users to provide any further identification or any further effort.

The features required may be difficult to legitimately collect, however. It is possible that software attempting to identify a user based on their typing habits may constitute key-logging which is in breach of many country's laws. It is also likely to suffer from similar issues to signature recognition (discussed above). A person's typing habits are liable to change based on a number of external factors such as fatigue or simply the task one is engaged in. A person's typing will also improve over time as they become more accustomed to typing on a particular keyboard and thus the latency's between key presses that the method above (and other methods such as Shepherd (1995)) focused on are liable to change.

2.6.6 Speaker Recognition

Speaker recognition focuses on identifying a user based on characteristics of their speech. Two main versions of speaker recognition exist: *text-dependent* and *text-independent*. Text-dependent speaker recognition attempts to identify a speaker based on a particular phrase, whereas text-independent speaker recognition attempts to identify the speaker without the need for a specific utterance (Microsoft 2006). In use cases where a user is already speaking, such as telephone banking, speaker recognition is incredibly useful and requires very little change in a user's behaviour. HSBC and other banks such as Citi have been using speaker recognition to identify their users for two years (Kollewe 2016). Text-independent speaker recognition is more resistant to fraud (Jain et al. 2004) but "requires longer training and testing utterances to achieve good performance" (Microsoft 2006). Given that text-dependent speaker recognition requires a specific phrase be said it could be combined with a password or PIN to enhance it's security.

3 Analysis and Discussion

3.1 Comparison and Contrast

Many surveys display a table like Table 1 to compare and evaluate different biometrics in relation to their performance, security and usability.

Biometric	**Fingerprint**	**Face**	**Hand Geometry**	**Iris**	**Voice**
Barriers to universality	Worn ridges; hand or finger impairment	None	Hand impairment	Visual impairment	Speech impairment
Distinctiveness	High	Low	Medium	High	Low
Permanence	High	Medium	Medium	High	Low
Collectibility	Medium	High	High	Medium	Medium
Performance	High	Low	Medium	High	Low
Acceptability	Medium	High	Medium	Low	High
Potential for circumvention	Low	High	Medium	Low	High

Table 1: Standard Biometric Comparison Table **From: Jain et al. (2003)**

This table can offer a good insight into which biometrics appear to be most effective, for example Fingerprints appear to be quite unique, very permenant (disregarding the case discussed in 2.2), fairly collectable and difficult to circumvent. This is likely why they are used in many consumer devices to add security. Compared to voice recognition it appears to be far better. Table 1 voice recognition is not very distinct, liable to change and quite easy to circumvent. The issue with table 1 is that it only contains a handful of biometrics, the author

(Jain et al. (2003)) has not evaluated many behavioural biometrics such as handwriting recognition, signature recognition and keystroke dynamics. It is likely that these were not available at the time the paper was written, however.

The main issue with table 1 is that it does not use numerical data to compare the different biometric methods. Most commonly when measuring biometric identification performance one measures "the percentage of queries in which the correct answer can be found in the top few matches" (Phillips et al. 2000). This differs to the measures used for verification, in this case one would measure the false-alarm rate and false-reject rate (Phillips et al. 2000). The false-alarm rate measures a system's likelihood to accept an invalid identity whereas the false-reject rate measures a system's liklihood of rejecting a valid signature. When developing a system targets for the false-alarm and false-reject rate can be set. These depend on the application and thus may make comparing implementations difficult as one false-reject/false-alarm rate that may be ideal for a particular system or implementation could not be optimal for a different system or implementation (Phillips et al. 2000). In contrast to these measures Tassabehji & Kamala (2012) uses the System Usability Scale (SUS) (Brooke et al. 1996) to evaluate the usability of a biometric system designed for use in online banking. The article evaluates the usability of different biometric systems finding that fingerprint scanning was deemed favourable amongst users in terms of ease of use and security. While usability of the biometric system is important it could be argued that it's perceived merits are secondary to it's actual performance as a secure identification method and therefore the false-alarm and false-reject rates discussed previously are more important.

Phillips et al. (2000) gives guidance on how researchers should evaluate biometric systems. The article suggests researchers should:

- Test on biometric signatures not yet seen by the system. The the idea being that to test on signatures the system has been trained with would simply result in assesing the ability to "tune a system to a particular data set".
- Publish the evaluation procedure, evaluation protocols, examples from the data set and performance results.
- Ensure the evaluation is not too hard or easy. If it is too easy or hard then the test will produce results that are not able to be compared to other tests.

When looking at evaluations of behavioural biometric systems many of the same measures that are used to evaluate biometric systems are presented. An example of this is Araujo et al. (2005) which uses false-reject rates and false-accept (false-alarm) rates to evaluate a keystroke-dynamics based system. The evaluation also appears to mostly follow the suggestions outlined in Phillips et al. (2000)

With the behavioural biometrics discussed in 2 & 2.6 it is clear that there are distinct use cases for each. With regards to the internet and identifying a person online the use of keystroke dynamics to augment an existing username & password system would enhance security whilst requiring very little extra effort on the part of the user. In cases where a user is already using their voice to interact with a system (smart speakers, telephone banking etc...)

it would again add further security to an existing PIN or pass-phrase based identification system without the user being required to do anything more than enrolment.

Whilst behavioural biometrics may not yet be secure and reliable enough to be used as a system's only form of identification they are certainly good candidates for use in multi-factor identification/authentication systems.

3.2 Evaluation

3.2.1 Comparison of Biometrics and Behavioural Biometrics

To compare biometric and behavioural biometric forms of identification and verification one should take into account not only the security but the impact on the user. With regards to the ethical implications of data collection it may impact users greatly if biometric features were accessed by a malicious third party. With reagards to the use of systems.

Both the biometric and behavioural biometric fields contain more and less secure methods, however behavioural biometric systems tend to be less obtrusive with how they collect data (Yampolskiy & Govindaraju 2008). While behavioural biometrics do not necessarily provide enough accuracy for identification, for verification purposes they appear to work well (Yampolskiy & Govindaraju 2008), and certainly would be well suited to multi-factor authentication systems.

4 Summary

The use of biometrics and behavioural biometrics for user identification is already becoming widely used in modern technology. Services like Apple's TouchID and FaceID help popularise biometric forms of identification and help keep user's data more secure that if they were simply authenticated with usernames and passwords or a PIN. Behavioural biometric identification methods such as handwriting recognition (2.6.3), signature recognition (2.6.4), keystroke dynamics (2.6.5) and speaker recognition (2.6.6) can all be used to enhance the security of existing systems without requiring much more effort on the part of the user. This is important as user's are likely to reject new technology that overcomplicates a system they already deem to be secure enough.

It is clear that usernames & passwords are not secure enough when it comes to securing user's sensitive data. Failings in the username & password system have been consistently highlighted through a number of high-profile data breaches and alternatives must be considered.

While biometric identification methods may have issues of their own, it can be proven that they are more secure than usernames & passwords and that they remove many of the issues that username & password identification faces, such as susceptibility to brute-force attacks or bad decisions made by users.

With the constant threat of security breaches it is imperative implementations of alternative identification methods be explored and considered. Username & password based authentication is proven to be greatly flawed and should not be solely relied on.

References

Abagnale, F. (2017).
URL: *https://www.youtube.com/watch?v=vsMydMDi3rI*

Ambalakat, P. (2005), Security of biometric authentication systems, *in* '21st Computer Science Seminar', p. 1.

Apple (2017), Face id security, Technical report, Apple.

Araujo, L. C. F., Sucupira, L. H. R., Lizarraga, M. G., Ling, L. L. & Yabu-Uti, J. B. T. (2005), 'User authentication through typing biometrics features', *IEEE Transactions on Signal Processing* **53**(2), 851–855.

Brooke, J. et al. (1996), 'Sus-a quick and dirty usability scale', *Usability evaluation in industry* **189**(194), 4–7.

Delac, K. & Grgic, M. (2004), A survey of biometric recognition methods, *in* '46th International Symposium Electronics in Marine', Vol. 46, pp. 16–18.

Eisen, O. (2017), 'Systems and methods for user identification using graphical barcode and payment card authentication read data'. US Patent App. 15/473,026.

Han, Y., Ryu, C., Moon, J., Kim, H. & Choi, H. (2004), A study on evaluating the uniqueness of fingerprints using statistical analysis, *in* 'International Conference on Information Security and Cryptology', Springer, pp. 467–477.

Jain, A. K., Prabhakar, S. & Pankanti, S. (2003), 'Biometric recognition: Security and privacy concerns', *IEEE Security & Privacy* **1**(2), 33–42.
URL: *doi.ieeecomputersociety.org/10.1109/MSECP.2003.1193209*

Jain, A. K., Ross, A. & Prabhakar, S. (2004), 'An introduction to biometric recognition', *IEEE Transactions on circuits and systems for video technology* **14**(1), 4–20.

Kessler, G. C. (1996), 'Passwords—strengths and weaknesses', *Disponível por www em http://www. hill. com/library/staffpubs/password. html (08 maio 1999)* .

Kollewe, J. (2016), 'Hsbc rolls out voice and touch id security for bank customers'.
URL: *https://www.theguardian.com/business/2016/feb/19/hsbc-rolls-out-voice-touch-id-security-bank-customers*

Microsoft (2006), 'Speaker verification: Text-dependent vs. text-independent'.
URL: *https://www.microsoft.com/en-us/research/project/speaker-verification-text-dependent-vs-text-independent/*

Ng, A. & MUSIL, S. (2017), 'Equifax data leak may affect nearly half the us population'.
URL: *https://www.cnet.com/news/equifax-data-leak-hits-nearly-half-of-the-us-population/*

Ogane, T. & Echizen, I. (2017), Biometric jammer: Preventing surreptitious fingerprint photography without inconveniencing users, *in* 'Biometrics (IJCB), 2017 IEEE International Joint Conference on', IEEE, pp. 253–260.

Phillips, P. J., Martin, A., Wilson, C. L. & Przybocki, M. (2000), 'An introduction evaluating biometric systems', *Computer* **33**(2), 56–63.

Shepherd, S. J. (1995), Continuous authentication by analysis of keyboard typing characteristics, *in* 'European Convention on Security and Detection, 1995.', pp. 111–114.

Tappert, C. C., Suen, C. Y. & Wakahara, T. (1990), 'The state of the art in online handwriting recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(8), 787–808.

Tassabehji, R. & Kamala, M. A. (2012), 'Evaluating biometrics for online banking: The case for usability', *International Journal of Information Management* **32**(5), 489 – 494.
URL: *http://www.sciencedirect.com/science/article/pii/S0268401212000898*

Yampolskiy, R. V. & Govindaraju, V. (2008), 'Behavioural biometrics: a survey and classification', *International Journal of Biometrics* **1**(1), 81–113.