Exploring Mobile Biometric Performance through Identification of Core Factors and Relationships

Matthew Boakes, Richard Guest, Farzin Deravi, and Barbara Corsetti

Abstract—Biometrics, as a form of authentication, has existed for several decades and shows no signs of slowing down. Extensive research has been carried out into enhancing systems either by improving error rates or ease of adoption by examining barriers to use. In this paper, we investigate factors of a biometric system that is likely to affect performance, in particular, focusing on mobile device implementation. By surveying the area, we have identified seven core factors that help to form a clearer understanding of what changes the performance of a system. These seven factors are Users, Modality, Environments, Diversity of Scenarios, System Constraints, Hardware and Algorithms and form 'The Core Factors Affecting Mobile Biometric Performance'. We utilise these factors to illustrate the practicalities of mobile implementations and indicate future considerations to explore future performance enhancements and provide an informative overview to developers, implementers and testers of biometrics systems, enabling the binning of performance alterations within one of these factors.

Index Terms—biometrics, performance, mobile implementation, users, modality, environments, scenarios, system, hardware, algorithms

1 Introduction

Biometric systems use automated methods to verify or identify an individual and have seen widespread deployment over the past two decades. Increasingly these technologies are being ubiquitously utilised on mobile platforms such as smartphones and tablets. Jain *et al*. [1] described seven factors to help assess the suitability of a human trait for biometric authentication one of which defines *performance* in that it "relates to the accuracy, speed, and robustness of technology used". There exists a vast field of research about performance claims for biometric systems; however these claims are usually predicated on a changed factor(s), and the resulting performance deterioration or improvement demonstrated. This paper aims to identify the core factors that need considering for the specification, evaluation and reporting of biometric systems.

Mansfield *et al*. [2] produced a comprehensive list of factors with the potential to affect the performance of a biometric system. Outlined in Table 1, these influencing factors were included within the ISO/IEC 19795-1:2006 [3] international standard on biometric performance testing and reporting including strategies for mitigation, such as a section on 'controlling factors that influence performance'. Although the strategies discussed in the standard maintain relevance, we must observe them with current developments in biometric technologies. One strategy in the standard suggests that "enrolment conditions should model the target application enrolment", but this is not so straightforward when we move from a static system to a mobile one where the enrolment conditions could be a plethora of different environments and scenarios. Furthermore, the standard only seeks to look at performance through very generalised metrics, mainly those of failure-to-enrol, failure-to-acquire, false match rate and corresponding false non-match rate.

In this paper, we intend to take an approach to separate the core factors that affect the performance of mobile biometric systems. There is excellent research [4], [5], [6] that discusses performance for traditional biometric systems, and these have usually been in the form of exploring *influential factors*. We wish to expand on these currently defined factors and identify new and more prominent areas that will need extra consideration when considering a mobile biometric system. We develop on how the community thought about 'Users' in the past and highlight the importance of usability when discussing performance we also illustrate how 'Environments' have a greater context when considering a mobile setting that links closely with the newly introduced factor of 'Scenarios'. Along with 'Scenarios', we establish 'Algorithm', and 'System Constraints' as new factors that have not previously been considered explicitly within a performance context. The Oxford Dictionary defines the term factor as "a circumstance, fact, or influence that contributes to a result" [7]. We see these seven factors as the fundamental factors of mobile biometric performance acting as a foundation layer for which more properties and areas can be discovered and connected but at the same time always linking back to one (or more) of these seven. The factors form unique connections with one another, meaning an impact on one can cause a performance alteration in another.

This study will allow developers to concentrate efforts more effectively when devising ways of testing and

- *Boakes, Guest and Deravi are all affiliated with the School of Engineering and Digital Arts, University of Kent, Canterbury, United Kingdom*
  M. Boakes E-mail: mjb228@kent.ac.uk
  R. Guest E-mail: R.M.Guest@kent.ac.uk
  F. Deravi E-mail: F.Deravi@kent.ac.uk
- *Corsetti affiliation is with the University Group for Identification Technologies (GUTI), University Carlos III of Madrid Leganes (Madrid), Spain*
  B. Corsetti E-mail: bcorsett@ing.uc3m.es

Manuscript received XX, X

analysing the performance of biometric systems in the future. Using existing research, we show the existence of each factor and demonstrate the impact on the overall performance.

While examining the definition of biometric system performance, it is necessary to include further input from Jain *et al*. [1] regarding *acceptability*; how well individuals "accept the technology such that they are willing to have their biometric trait captured and assessed" and *circumvention* which "relates to the ease with which a trait might be imitated using an artefact or substitute".

Both of these issues we feel are critical when considering the overall performance of a biometric system. We believe acceptability is essential because, within the technology sector, public perception and acceptance will be one of the reasons that prevent the uptake of a biometric system's use. A recent example of driving public opinion by privacy concerns happen this year in San Francisco, where public administration has recently banned the use of facial recognition for local services [8]. Although this is a rather extreme example, it does highlight that it is necessary to consider the user's acceptance for a system before rushing ahead with the implementation; otherwise, it will alienate the users. Circumvention primarily refers to spoofing the biometric system using artificial means. Any system that offers a biometric solution will require to have some level of resistance to these attacks. If a system allows a large amount of fake or artificial produced modalities as a genuine user, we believe this highlights major flaws and hence a massive decrease in performance.

By considering *acceptability* and *circumvention* we are challenging the conventional approach and definition of *performance*, we focus with the end-users on how they perceive the use of the system and how both *good* and *bad* actors may attempt to trick the system using various means, we believe these are vital areas when considering the performance of the overall system. The metrics discussed in ISO/IEC 19795-1:2006 [3] also ignore the usability aspect, which we illustrate to be of vital importance to the performance of a system.

Section 2 of this paper will briefly discuss the importance that mobile biometrics have had on producing this work while Section 3 will provide an insight into our methods. Sections 4 – 10 will discuss each influencing factor in more detail. Section 11 will introduce the relationships between them and Section 12 will begin to provide a practical example of use. The final Sections 13 – 14 will provide conclusion and future work considerations.

2 Mobile Biometrics

A central reason for the need to update the current understanding of 'performance' for biometric systems is the proliferation of mobile biometrics. Throughout this paper, we will refer to mobile biometrics concerning smartphone devices. Although the traditional uses of biometrics remain (such as border control systems and national ID cards), there has been a widespread adoption into the mass market brought about due to the incorporation of specific biometric sensors into current smartphone devices.

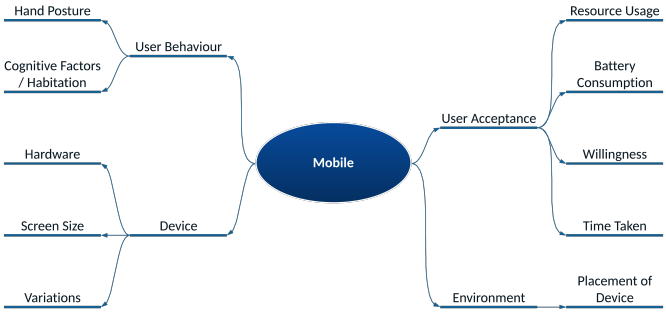

Fig. 1. Examples of Mobile Influences Expected to Impact Performance

Mobile biometrics provides a multitude of novel opportunities to explore and offer a convenient and arguably more secure way of providing authentication. An example of the adoption of this technology is with mobile banking to enable payments through services such as Apple Pay and Google Pay. Applications on these devices can make use of the embedded biometric technology to provide convenient access to their services without the need to enter and remember a password each time or to access personal information.

This adoption of biometrics to a mobile market requires us to rethink how we go about the approach to testing and verifying that the system is *fit for purpose*. Moving from the more traditional fixed (static) system to a mobile (dynamic) one increases the environments and scenarios in which they will operate. In turn, this has a knock-on effect on how the users will perceive and use the system.

Figure 1 begins to form an initial collection of performance influences for a mobile context. We assembled these influences through reading the literature [9], [10] and assessing what factors are introduced or more commonplace when a mobile scenario, such as on a smartphone, is used. Factors such as screen size and user posture can influence usability and have an impact when it comes to touch-based mobile biometrics [11].

Research studies are beginning to explore these factors [12], [10]. However, there are still limited resources on how the performance of mobile biometrics change in a variety of scenarios and environments requiring further exploration, including the definition of a suitable methodology. We have drawn on evidence from existing conventional systems, including signature analysis, to begin to plug the gaps in where these limited research resources currently are.

3 Methodology

The motivation for this was to look through a user's perspective and begin to form a model that would ultimately allow for the specification, evaluation and reporting of a mobile biometric system. The output was to produce a state-of-the-art with considerations for how we can provide core factors that will enable binning categories for identifying the reasoning behind changes in biometric performance. Initially searching for articles using search terms linked with '[Mobile] Biometric System(s) Performance' led to journal databases and library catalogues and following chains of

TABLE 1
List of Influencing Factors as Defined in ISO/IEC 19795-1:2006 [3]

Factor	**Description** [2], [3]
Population Demographics	Characteristics of a population, such as the age, gender, ethnic origin and occupation
Application	The overall system itself, such as enrolment and verification elapsed time, user familiarity and user motivation
User Physiology	Physical properties of a person, such as beards, skin tone, height and disability
User Behaviour	Behavioural properties of a person, such as a dialect, movement, stress and facial expressions
User Appearance	How a person looks, such as clothing, hairstyle, bandages and tattoos
Environmental Influences	Factors of the environment, such as background, lighting level, weather and reflections
Sensor and Hardware	The factors affecting the devices correct operation, such as dirt, focus, sensor quality and transmission channel
User Interface	Means by which the user and a computer system interact, such as feedback, instruction and supervision

references introduced new material. While exploring the literature, we were able to begin to categorise them into the performance factors investigated in each paper. This approach led us to the discovery of the core factors, and the literature most relevant to this discussion. It quickly became apparent that having multiple core factors would allow for separating the literature into suitable groupings. Once we saw these core factors forming, we were able to look more explicitly using them as additional search terms to find literature that helped to support the model. During the survey, some research papers could categorise into more than one category due to more than one factor being present in the article. We include literature that investigated an apparent performance factor(s) where the impact was evident and discarded any that were too broad in approach and were irrelevant to the discussion made here.

TABLE 2
Examples of Influencing Factors per Modality

Modality	**Sample of Influencing Factors** [13]	
Face	• Movement	• Facial Expression
	• Age	• Skin Tone
Fingerprint	• Fingerprint Condition	• Weather
	• Arthritis	• Offsets and Rotations
Voice	• Ethnic Origin	• Noises
	• Colds or Laryngitis	• Misspoken Phrases
Iris	• Lightning Level	• Eyelashes
	• Blindness	• Reflections
Signature	• Age	• Injuries
	• Sensor Pressure	• Motivation

4 Factor #I: Modalities

With existing biometric modalities, we can target the strategy to consider known influencing factors that are likely to affect the performance and include these within the testing strategy. Known as '*influencing factors*' Mansfield *et al*. [2] provided a list detailing a significant amount of these factors, detailed in Table 1. Each modality directly introduces a set of influencing factors caused by choosing that particular modality over others, like how wearing glasses will likely affect an iris recognition system.

Table 2 shows illustrative examples of influencing factors for some common biometrics that we have identified and is by no means a completely comprehensive list. It can be seen from these factors in both Table 1 and 2 that the users are a big influence in these factors that affect the modality.

The modality is going to be uniquely linked to the remaining factors, for example, the hardware and algorithms, as each will have a performance impact that could affect the other. It is also essential to know the *influencing factors* that are specific to each modality when used in particular scenarios as this will help to identify the areas requiring increased attention when testing a biometric system. For example, appearance is going to have little impact on voice but is likely to have a more significant effect on a facial recognition system.

The understanding that a single modality is not without its issues is apparent as researchers have made significant advances with multimodal biometric systems, intending to combine separate modality systems in a way that allows for an improvement in recognition accuracy performance. Part of these multimodal systems tries to overcome the '*influencing factors*' issues through the combination of the results of another trait that will hopefully not be affected, and hence reduce the error that would otherwise have happened [14].

Jain *et al*. [14] reported that unimodal systems (using a single trait/modality) experience several problems including "noisy sensor data, non-universality and lack of distinctiveness of the biometric trait, unacceptable error rates, and spoof attacks" all of which can affect the overall performance of a system. In comparison, He *et al*. [15] explored the performance of a multimodal system using three traits: fingerprint, face, and finger vein. The results of the experiments concluded that "multimodal biometric system can achieve significantly better performance compared to a single biometric system" but also that the inclusion of adding finger vein "results in a verification system with very high accuracy". This research demonstrates how the performance changes when using a variety of biometric traits which confirms that performance can be affected by modalities.

Gafurov *et al*. [16] assessed the use of gait as a biometric trait and concluded that it is better suited as a "complementary biometric" and not as a "replacement for traditional authentication mechanisms". The work also noted that there are "several factors that may negatively influence the accuracy". They classed the factors for gait as

'*External*' (viewing angles, lighting conditions) and '*Internal*' (sickness, physiological changes). They identified how gait was "robust against minimal effort impersonation attacks" and concluded by noting that an "investigation of these factors is very important towards developing robust systems" which identifies how necessary it is to appropriately select a modality for a particular scenario in a way that will try to mitigate issues (caused by influencing factors).

Ito *et al*. [17] commented on how researchers seek "new biometric traits to enhance the accuracy and convenience of biometric recognition" suggesting that modalities differ in their performance. Each modality is unique and with that brings an assortment of '*influencing factors*' to contend with, although some of these factors are common between specific traits. It is undoubtedly true that a modality holds extraordinary power over the system in that it can define how the rest of the system develops around it, meaning it is a priority to select the right modality for the job. An example of a lesser researched biometric includes that of the 3D ear shape, which was explored by Yan *et al*. [18].

Furthermore, the user in terms of sample quality as well as age (elderly) [19] and accessibility (wheelchair users) [20] will influence the performance of the modality. Elliott *et al*. [21] investigated how fingerprint sample quality across age groups can affect a biometric system. They concluded that "more emphasis should be placed on an individual's age, rather than the moisture of the finger when developing a fingerprint recognition system" as the quality of the image became more variable for an older population (aged 62 and over).

Due to the '*influencing factors*' present, the modality themselves have an impact on the performance of a system. It is for this reason, therefore that it is considered to be one of the factors.

5 Factor #II: Environments

The environment can significantly impact the performance of a biometric system. These systems are becoming less fixed and more mobile, and the environments in which they will operate are becoming nearly impossible to predict.

Research produced by Lunerti *et al*. [22] examined the impact that environmental factors had on face recognition on smartphones. They assessed facial image quality (FIQ) in both indoor and outdoor conditions and was able to conclude that "[biometric] scores obtained with the images taken from the smartphone are higher with the images taken indoor" showing that the environment has an impact on the performance of a system.

One of the main aspects that allows a mobile biometric system to differentiate itself from a traditional biometric system is the sheer range of environments and conditions that the device will be required to operate within. The performance could be affected at both the enrolment and authentication phase due to the variety of these different environments and situations that could occur while the process is happening, meaning that a robust enrolment template, captured under 'optimal' conditions, may not produce accurate matches with samples collected under certain circumstances at later verification attempts. Furthermore, an enrolment template captured under poor conditions may not be functional at all.

TABLE 3
Examples of Scenarios under Categories of 'Motion' and 'Stationary'

	Motion		Stationary
User	**Transportation**	**Dual**	
Walking	Bus	Walking on a train	Sitting
Running	Train	Walking on a boat	Standing
Cycling	Earthquake	Swimming	Lying Down

Different environments that could have an impact on the performance include:

- Indoors vs Outdoors
- Lighting
- Weather Conditions
- Terrain – physical features of the environment, from the ground being walked over to the type of location (e.g. city, countryside, ocean)

Previously Elliott *et al*. [21] had also shown how illumination could have a significant effect on the performance of facial recognition systems. They concluded that "enrolment illumination level is a better indicator of performance than the illumination level of the verification attempts" and found that the "enrolment light level should be as high as possible" when the "lighting conditions are not constant for verification".

With regards to behavioural biometrics, the performance of voice recognition severely degrades when ambient noise is present, as shown by Gong [23]. Research has been carried out to mitigate and detect this noise and hence the performance of a voice recognition system. Yamada *et al*. [24] "described a method for estimating the performance of a speech recognition system using a distortion measure".

Applying the environment concept to conventional modalities has been known to affect performance. These conditions include:

- **Face** - Background (Multiple Faces)
- **Fingerprint** - Weather
- **Voice** - Noise
- **Iris** - Illumination

These studies show how the environment comes to impact biometric performance, and we deem it worthy as another factor.

6 Factor #III: Diversity of Scenarios

The 'Diversity of Scenarios' relates to how an individual is using a device within an environment. We believe these scenarios can be classed under two categorise using the headings 'motion' and 'stationary'. Table 3 shows a brief table of example scenarios.

In our classification, 'motion' refers to the scenario of being in movement relative to the environment and, likewise, 'stationary' being where the device is at rest (no action) corresponding to the environment. Furthermore, we define 'transportation' as any scenario where the device is

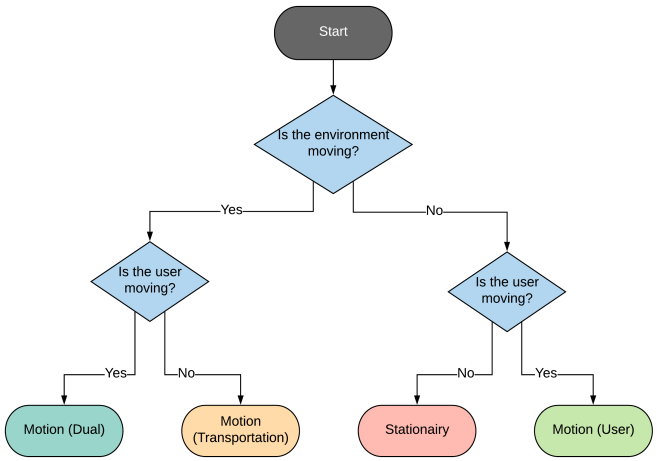

Fig. 2. Flowchart to assign a scenario to a category of 'Motion' or 'Stationary'

in motion, and the cause is an external influence, from the environment, not coming from the user (e.g., when being driven around such as on a bus). This splitting of 'motion' scenarios into 'user' and 'environment' concepts introduces an overlap scenario where the cause is a combination of both a user and environmental factor, and we define this as a 'dual' motion scenario. Figure 2 illustrates a flowchart to aid in assigning scenarios to a potential category. We overlook subtle minor movements that will likely occur in all stationary situations such as shaky user hands while holding the device.

There is currently limited research assessing the variety of scenarios under which a biometric authentication can occur, including for the conventional modalities: Face, Fingerprint, Voice and Iris. However, different scenarios can alter the performance of a biometric system. For example, Blanco-Gonzalo *et al*. [10] explored performance changes across conditions when signing using dynamic signature verification (DSV) systems. Their results showed, although there is "not an ideal scenario for signing", performance improvement is observed when using a stylus device with "the user sat on a chair, and the device is resting on a table" and for finger-stylus devices when "the user has to handle the device without support".

The location of a scenario plays a role in affecting the user's behaviour and state of mind, for example, "stress influences negatively both performance and usability" [10]. Consider the scenario of signing within an outlet such as a post office. Here the scenario encourages users to sign quickly and carelessly to avoid causing long delays, which can introduce stress and anxiety and can affect the performance negatively. Whereas, in a ceremony-based scenario, such as the signing of a legal document, the "user typically signs with greater care, striving for enhanced quality and clarity", demonstrated in research from Guest *et al*. [25], which causes an increase in performance.

The scenario links closely with the interaction the user has with a system in a particular environment and how adjustments may need to be made to account for these changes. These adjustments can come both from the system itself or from the way the user interacts with the system. The development of the Human-Biometric Sensor Interaction (HBSI) model [26] investigated how scenarios can modify performance. Brockly *et al*. [27] concluded "the development [of HBSI] reveals the complexity of the potential interactions and the changes of those interactions when digitisers change, as well as when the ceremony changes".

The number of scenarios that biometric authentication can occur increases dramatically when introduced to a mobile environment. Bhagavatula *et al*. [28] assessed the usability of a range of mobile biometrics systems. Firstly (and probably the least surprising) was that the Android "face unlock was completely unusable in a dark room". They also explored Apple's Touch ID and found that the Touch ID and face unlock "mechanisms fail in specific scenarios, wet fingers and dark rooms, respectively". They also conducted a series of walking experiments, one merely walking and another walking while carrying a bag in one hand, this was performed in laboratory conditions (indoors). They found that participants "did not find unlocking to be difficult for any authentication scheme in either of the walking scenarios". The experiment was mostly conducted from a usability perspective, concluding that participants preferred the Android face unlock in the walking scenarios as they were able to handle the phones in their desired positions. Whereas, for Apple's Touch ID, they had to hold the phone more precariously from the bottom as this is the location of the fingerprint sensor.

Sitová *et al*. [29] introduced "hand movement, orientation, and grasp (HMOG)" to authenticate smartphone users continuously and their work investigated two conditions, sitting and walking. The results showed that "HMOG improves the performance of taps and keystroke dynamic features, especially during walking" they theorised that this improvement was "attributed to (a) the distinctiveness in hand movements caused by tap activity and (b) the distinctiveness in movements caused by walking".

With increased development in mobile technology, research is looking into ways to help capture and authenticate a biometric trait while in motion, including work to perform long-range iris recognition, also known as iris-on-the-move, as surveyed by Nguyen *et al*. [30].

Given the proliferation of biometrics on mobile systems, this indicates that we should conduct further work in this area, and, for this reason, we believe it should be considered a factor.

7 Factor #IV: Users

Users of a biometric system need to have confidence in the authentication process and to gain this confidence; biometric systems need to be universal (applicable to all) for the end-users of the system. We choose modalities due to the uniqueness they provide in being able to identify between individuals. However, being able to develop a system to accurately extract all the small features within a modality to achieve the uniqueness is not an easy task, and this is how false positives can occur.

One of the most prominent ideas in biometric testing is the concept of the '*Biometric Zoo*' proposed by Doddington *et al*. [31]. In this work, they used voice recognition to show

that users of a system could directly have an impact on the overall performance. This work proved the existence of different categories of users:

- **Sheep** - Sheep dominate the population and systems perform nominally well for them.
- **Goats** - Goats are those speakers who are particularly difficult to recognise.
- **Lambs** - Lambs are those speakers who are particularly easy to imitate.
- **Wolves** - Wolves are those speakers who are particularly successful at imitating others.

They state that "goats have the greatest performance effect" adding a considerable amount of false-negative data, whereas the wolves and lambs attribute more to the false-positive data and hence end up affecting the overall performance. The biometric zoo or menagerie was extended by Yager *et al*. [32] to include further groups of users that cover the extreme ends of the spectrum and explores the existence of the menagerie within other modalities:

- **Worms** - Worms are the worst conceivable users and match poorly against themselves.
- **Chameleons** - Chameleons always appear similar to others and receive high match scores.
- **Phantoms** - Phantoms always receive low match scores regardless of the comparison template.
- **Doves** - Doves are the best possible users, matching well against themselves and poorly against others.

The framework presented here highlights weaknesses in the system and, should any of these user groups exist, whether that be within the algorithm itself, the enrolment quality or data integrity. They conclude by saying that the "biometric menagerie is a diagnostic tool that takes a more user-centric approach".

The biometric menagerie is not without its critics Popescu-Bodorin *et al*. [33] claim the concept is 'fuzzy' as to whether the categories are referring to the users themselves or the templates. Part of their claim highlights that the category of the users can change based upon the calibration of the system. Although this may be true, the concept of the biometric menagerie is still one that we believe is useful for highlighting how users can affect the performance or how users can be used to identify potential flaws and weakness within a system. In either of these cases, the users directly affect the performance.

The user interaction with the interface and user-acceptance of the modality and scenario is a factor of performance which is often not considered. However, if users encounter a bad experience in using the system, it may result in an unwillingness to use the technology on an ongoing basis. As well as examining the environment impact, Lunerti *et al*. [22] also examined the effect of user-interaction with face recognition on smartphones and found, through a questionnaire given to the participants, that the ease and confidence users had with the system increased with each session when operated indoors, however, when used outdoors the confidence remained relatively constant throughout.

As noted previously, research has also shown that the physiology of a user can affect the performance of a biometric system, including age [21] and accessibility [34].

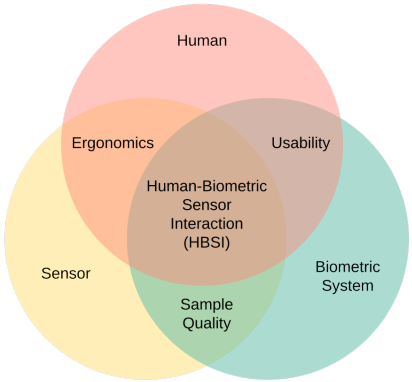

Fig. 3. Human Biometric Sensor Interaction (HBSI) Model [38]

Another important factor when discussing how users affect system performance is to examine users' acceptance and satisfaction of biometric systems as this, in turn, will not only drive the overall performance but will indicate the willingness to use such a system. El-Abed *et al*. [35] proposed that "taking into account users' view . . . is not only beneficial to the end-users, but it will also help to improve performance and effectiveness". Thoughts and opinions from users can be used to influence the design and interface of the biometric system. Respondents of a survey conducted by El-Abed *et al*. found that "biometric-based technology is more appropriate than secret-based solutions against fraud" and that the "trust factor has been identified as a major [aspect] that affects their general appreciation". It is also worth noting that the user's culture can influence the acceptance and use of biometrics [36], [37].

By examining this research, it has become clear that users' acceptance and willingness are crucial factors to consider investigating when it comes to using a particular biometric system. It is possible to imagine a scenario where a system has excellent performance results; however, this does not provide any guarantee that the users will be inclined to engage with the system. The affect usability has on performance is becoming an increasingly relevant research topic as traditional performance metrics have relied only on error rates, which generally do not consider usability concerns.

Shown in Figure 3 is a thematic outline of the Human-Biometric Sensor Interaction (HBSI) model aimed at addressing this issue. Guest *et al*. [25] identified that "performance deterioration of a tuned biometric software system might be caused by an interaction error with a biometric capture device". They also discovered that when a biometric system is "deployed within a public setting . . . performance of a system drops, not because of a change in the algorithmic implementation". This discovery points to the need to include *Algorithms* as a separate factor which can affect performance. Finally, they concluded that in the case of dynamic signature verification "these problems can be solved through the design of appropriate on-screen user interfaces and hardware" which strengthens the argument for having *System Constraints* and *Hardware* as separate factors that can affect performance.

Miguel-Hurtado *et al*. [39] assessed Voice using the HBSI

model for a mobile authentication system on smartphones. Their results concluded that "the learnability of the application needs to be improved by better guidance . . . thus, better user interfaces and participant guidance within the application have been recommended". They noted that this would improve the overall performance by "avoid[ing] user's assistance requests and reduce the user's errors. Hence, it will help to reduce the number of incorrect presentations and raise the rate of successful enrolments".

Jain *et al*. [1] made a comparison of biometric traits using data of the perception of three biometric experts. This comparison showed the acceptability given to the common modalities, on a high, medium and low scale as follows:

- **Face** - High
- **Fingerprint** - Medium
- **Voice** - High
- **Iris** - Low

As mentioned we believe acceptability is essential but the concept may have grown into general public perception along with privacy and security concerns. Acceptability forms one of the reasons that will prevent the uptake of a biometric system's use and highlights the necessity to consider the user's acceptance before rushing ahead with the implementation.

When assessing the performance of mobile biometrics, it will be necessary to identify the familiarity a user has with the device in question. The reason for this is regarding habitation. Users who are more familiar and comfortable with a device are likely to perform better than someone who is handling it for the first time. We suggest that all users have time to adjust and familiarise themselves with the device before any formal testing begins to mitigate any performance impact from different levels of habitation from the users involved.

Users are a significant factor that affects the performance of biometric systems, so it is little surprise to include them within the core set.

8 Factor #V: System Constraints

It is necessary to consider that systems often need to meet their own needs and demands for the scenario in which they required to operate. These needs will include requirements such as:

- Verification or Identification?
- Throughput Rate
- Required Error Rates
- On-Device vs Off-Device
- Time to Enrol or Authenticate
- Privacy Protection / Control
- Latency

On-device refers to the processing of the biometric algorithm occurring on the hardware of the device of the biometric system, whereas off-device is the scenario where some or all the processing gets delegated to external equipment for processing, most likely a server. Each introduces security concerns that will need exhaustive testing and ties into the privacy protection and control provided.

Privacy protection and control refers to the security provided for storing all captured samples even when the authentication algorithm is performing the comparison between samples. Securing these templates is crucial to gain user confidence and ensure the system is not vulnerable to attacks. However, there will be a trade-off between security and performance as more significant restrictions will generally mean slower functionality. Latency is the delay introduced by transferring data around, and this could be an issue for off-device systems where the captured sampled needs to be transmitted for processing or in providing visual cues to users.

These specific requirements will introduce their constraints and will impact the performance of the system. For example, whether authentication is to take place off-device, most likely on a server, we must note how the performance is affected under different network setups (including Wi-Fi, 4G, and 3G).

A considerable number of factors within *System Constraints* link closely to that of overall user experience. A system will likely function as intended but, due to the constraints placed on the system by its requirements, it now takes a long time to perform a verification then its performance is going to suffer. A *System Constraints* can be defined as something that introduces a constraint on how the system can function, this could be due to many factors, including external (corporate) restrictions, device requirements and scenarios.

Users enjoy convenience and ease, and hopefully, when implemented correctly, this is something that biometrics can offer as a service. A classic use-case for biometrics is in airport security as a way to process large numbers of people as quickly and efficiently as possible. Sasse *et al*. [40] investigated a range of biometric process at various airports. This scenario contains a lot of constraints and requirements that any biometric system will need to adhere to, thus creating *System Constraints*. The study mentions how a significant factor here is how the users react to the system and that any implementation should "emphasise usability's importance in successfully operating biometric systems".

Universal access is also one of the primary requirements associated with an airport border control system. Early tests showed how disabled users struggled to both be enrolled and later verified with the system. The other issue was the experience of the "*bendy shuffle*" as Sasse *et al*. defined the scenario when trying to position the body correctly for the verification sensor. This scenario was due to the interaction being entirely different from that of the enrolment phase caused by the fixed position of the sensors. There is a hardware issue present here. However, a verification process that was different from the enrolment has equally caused problems, and this was a flawed design of the system.

While exploring recent advancements in biometric recognition, Ito *et al*. [17] states, "biometric techniques [that are] to be used in the practical system depend heavily on application requirements".

Research has also explored the use of visual feedback and how this affects performance. Visual feedback is where visual cues help guide the user through the biometric process and provides suggestive feedback. As expected, "the better [the] visual feedback, the better performance and

usability" was demonstrated by Blanco-Gonzalo *et al*. [10]. They also showed that "users do not feel comfortable when [no] visual feedback is provided". While exploring dynamic signature verification (DSV), Blanco-Gonzalo *et al*. found that "latency ... involves annoyance in users, and it also affects the performance". Here the latency refers explicitly to the digital ink appearing on display providing a visual aid to the user.

However, latency does raise questions towards future concerns regarding the performance of systems that require off-device processing and how network latency can cope with the movement of data, thereby affecting performance.

Exploiting all the available resources in a system, besides the dedicated biometric sensor, may result in an improved recognition system. Using a smartphone example, where the use of all available sensors and hardware can be used to provide a continuous authentication mechanism, such as with the available touchscreen gestures as explored by Feng *et al*. [41]. When observing keystroke dynamics on a mobile device, Buschek *et al*. [42] were able to "improve implicit authentication accuracy through new features" available on a smartphone. They were also able to "improve usability with a framework to handle changing hand postures".

We also feel it is again necessary to mention *circumvention* from Jain *et al*. [1]. Here measures will explicitly need to be incorporated into a system to prevent security concerns and vulnerabilities. There will likely be the introduction of trade-offs between having a secure system and error rates along with the time to enrol/authenticate. All of which will mean balancing the performance. It is for these reasons that *System Constraints* belongs as a factor.

9 Factor #VI: Hardware

A biometric system can only be as good as the hardware it has to function on. This statement can be taken both in the sense of speed and functionality of processing and in the resources available to be exploited. The sensors, both dedicated biometric or otherwise can affect the overall performance of the system. Jain *et al*. [43] explores the performance of smartphone touchscreens with the traditional hardware keyboards, using the same modality of keystroke dynamics. Owning to the fact that touchscreen sensors "provide considerably richer data" they were able to exploit this data to generate results that demonstrated that "touchscreen data has considerably greater biometric value than that available on hardware keyboards".

Obtaining a detailed hardware description is useful in gaining a more detailed understating of performance, especially when considering authentication involving multimodal biometrics where one sensor may capture multiple modalities or where various sensors are used to each capture a single trait.

Elliott *et al*. [21] noted that while exploring signature capture systems "various devices used in studies demonstrate different physical and measurement characteristics". This message is still relevant today within a mobile context which provides a more excellent range of hardware devices on which we can implement biometric authentication. Developing a testing framework for this purpose will involve trying to take considerations for the variability between devices to ensure consistency.

"Hardware properties can affect the variables collected in the data acquisition process, and therefore the quality and performance of the device" [21]. How available sensors collect the acquisition features will also affect the performance of a system. Elliott *et al*. again concluded that "there are significant differences in the variables across devices, yet these variables are not significantly different within device families".

The introduction of biometrics into the mobile market is a relatively recent event, with the first smartphone to feature the biometric technology occurring in 2004 [44]. However, significant adoption was due to the introduction of Touch ID into the Apple iPhone series [45].

Hwang *et al*. [46] explored the implications of *portable biometric authentications*. Although this research is from 2004, it is interesting to see how the problems and experiences are still relevant today. They stated that a potential scenario is "financial and commercial transactions as a replacement for (biometric) smart cards" - we now observe that a significant end-use for biometrics on smartphones is the introduction of mobile payments including Apple Pay and Google Pay. Factors also discussed such as where to store the biometric template within the system (particular when portable) are crucial in the design and overall performance of the final product, and this is one of the critical arguments for the trade-offs surrounding on-device vs off-device:

"*Performing the biometric processing on the server provides performance benefits with significant security problems. Performing all the biometric processing locally [on the device] provides the best security, but requires a relatively larger amount of energy and latency*" [46].

We can observe this statement when examining a system's hardware along with the limitations and benefits it provides. It is hard to estimate whether people expected the current advancements in computing and particularly how powerful are smartphones have developed. Moore's law, the observation that "Manufacturers ... [have] been doubling the density of components per integrated circuit at regular intervals" (every two years) as surveyed by Schaller *et al*. [47], has been used as a reliable method for calculating and predicting future trends. Simplistically it is the application of this law which has allowed for higher-performance computers. The same is true for smartphones and the mobile market; with a continuing drive from industry exploiting hardware resources, current predictions show that Moore's Law is still likely to be accurate until around 2050.

As stated previously it can be challenging to predict the future outcome of the computing industry, but it is clear that within the realm of biometric systems the hardware used to support the system will have an impact on the performance for the system. Cantó-Navarro *et al*. [48] developed a floating-point accelerator "specially designed for accelerating biometric recognition algorithms" for embedded systems. They achieved this by exploring ways of accelerating the stages that usually proved the most time-consuming for biometric systems, such as Support Vector Machines, Gaussian Mixture Models and Dynamic Time Warping. They were able to obtain "acceleration factors

ranging from x7 to x22" on two complete biometric algorithms.

Emerging developments in cloud computing and mobile systems have shown that by effectively using the cloud, processing can be moved from a mobile device to save on the demand for power consumption and storage capacity. Smartphones today are currently capable of storing and running a biometric system without the need to offload resources. However, the same is not valid for mobile IoT devices where resources are limited.

Hu *et al*. [49] explored the use of cloud computing and Internet of Things (IoT) to create a functioning biometric system as IoT device typically do not have the same level of processing and storage capacity as the modern smartphones. They were able to create a face identification system that could "meet the growing demands of computation power and storage capacity in current big data era" by utilising the advantages of cloud computing with the parallel resolution mechanism. "In this scheme, resolution services and identity information management services are deployed in the cloud which can make full use of the high reliability, high scalability, powerful computing and storage capacity of cloud computing to provide efficient and accurate face resolution services". They admitted that the system was not without drawbacks, which included storing templates in a third-party data centre and the privacy and security associated with the overall system.

The captured sample quality produced from the biometric sensor needs to be of a high enough quality to be able to satisfy the biometric system. Poor quality images may result in more false rejects and cause the performance to suffer as a result. Metrics exist to be able to measure sample quality, including Face Image Quality (FIQ) and NIST Fingerprint Image Quality (NFIQ) algorithm.

It is for all these reasons that we wholeheartedly believe hardware is a factor affecting the performance of biometric systems.

10 Factor #VII: Algorithms

Algorithms are the computational backbone of a biometric system. Recent advances in machine / deep learning have allowed for significant progress in the field of computer vision as well as speech recognition, natural language processing and many more. It has therefore also found its way into biometric authentication. Conventional machine learning methods including Support Vector Machines, Principal Component Analysis and Linear Discriminant Analysis have provided the backbone algorithm for biometric systems in the past, but now with more 'deep learning' approaches discovered, it is likely that a rise in performance will occur as we produce more accurate machine learning models.

Taigman *et al*. [50] have experimented with deep learning on a 3D face model and have developed a system called 'DeepFace' which claims to "reduc[e] the error of the current state of the art by more than 27%".

Examining the conventional algorithms, He *et al*. [15] explored the performance comparison of sum rule-based score level fusion and support vector machines (SVM)-based score level fusion for multimodal systems and discovered that SVM "could attain better performance . . . provided that the kernel and its parameters [were] carefully selected".

It is clear that an algorithm and hardware are closely linked where the algorithm has to be able to perform on the available device to ensure the performance is usable, this involves making careful considerations for the amount of memory available to run the algorithm without causing a significant time delay for the users. Cantò-Navarro *et al*. [48] proved this as they were able to achieve higher acceleration performance, regarding execution time, by altering the hardware components to be more efficient for a biometric system.

Algorithms are being designed to produce higher accuracy results and overcome certain environmental factors. Face image processing is a significant research topic and covers many fields including computer vision, pattern recognition, image processing and biometrics as surveyed by Ito *et al*. [17] in which they state, "a variety of face image processing methods has been proposed since the performance of face image processing is significantly influenced by environmental changes such as head pose, expression and illumination changes".

Mobile biometrics provide another challenge for the algorithms being developed for those devices as they will need to contend with an array of unconstrained conditions to maintain a high level of operation. Biometrics is an exercise in pattern recognition, and machine learning algorithms have proved to be extremely useful in this area. Similarly, it has shown that "machine learning offers several advantages over other approaches for biometric pattern recognition" as discussed by Ortiz *et al*. [51], while also "satisfy[ing] an increasing need for security and smarter applications". Similarly, Blanco-Gonzalo *et al*. [10] stated that "the objective of the algorithm is to decide whether the user is the one who claims to be or not". With this in mind, the whole system's functionality depends on the algorithm, and it is for this reason that it is a factor.

11 Modelling Factor Relationships

Figure 4 shows the interaction model between the factors. These relationships (connections/links) show an association between factors, and we see them forming in several ways, such as being constraints or having an effect on the behaviour of each other.

We see many connections demonstrating the view that an alteration in performance may be caused by several of the factors discussed here. It is entirely plausible that an adjustment in one of these factors could incur a knock-on effect to another, for example, should the hardware be modified this could cause the functionality of the algorithm to change producing a poor implementation for feature extraction causing more false positives to occur. Similarly, a relationship may connect more than two nodes.

Figure 4 is an interpretation of where we see connections forming; however, that is not to say that this is a definitive model and more relationships may exist. The model is our first attempt at forming relationships between our factors, and it is not a comprehensive list. There will be relationships (links) missing, and we encourage others to find links and

continue to use and adapt our model in the attempt at forming a complete model. The model here begins to present the metrics for reporting the performance of a mobile biometric system, with the connections being some of the key features that a report should include to provide the assurance users need. Definitions of the relationships (mostly from the Oxford Dictionary [52]) is provided in Table 4.

The connection model highlighted *Users* as one of the most influential factors. An important factor in determining the performance of a system we can attribute to user satisfaction, however "the users' satisfaction is most of the times put aside" as highlighted by Blanco-Gonzalo *et al*. [10]. However, its importance is evident as "a non-usable system has not only repercussions in performance but users' acceptance of the technology also".

If we can find of way of qualitatively or quantitatively defining these relationships, then we hold a firm belief that we will be able to generate a value that can universally express performance. Some of the identified relationships already have robust research methods for obtaining a quantitative value such as retrieving sample qualities of biometric traits as discussed in current ISO standards [54]. However, the same is not valid for all of the relationships identified here and gathering all this information would be impractical for testing one system and we, therefore, propose that a subset of the data will be sufficient. We theorise that by treating the measure of each relationship separately, we can begin to build up an overall picture of performance and become more confident in its value with each newly added piece of information. The aim would be to allow us to compare different devices more consistently with one another, but the practicalities of this will require further research.

Table 5 begins to from suggestions about how we can start to collect the relationship data defined here along with some suggested basic examples. This table is by no means a complete list, but it is currently the formation of some initial ideas that will require updating with further analysis of best methods and practises.

12 Trialling the Model

To enable us to begin to test and apply the modelling and methodology, we apply this to a current high-end device, the Samsung Galaxy S9. This device allows users to enrol three modalities, Fingerprint, Face and Iris.

We present here the preliminary results of a recent data collection. A total of 60 users enrolled the three modalities while seated and holding the device comfortably in their hands. The users were then required to authenticate themselves in a variety of scenarios for a minimum of five transaction attempts. We present four of the scenarios, while the user is 'Sitting', 'Standing', walking on a 'Treadmill' (at a personalised speed) and walking down a 'Corridor'. Along with the four scenarios, we analysed and a 'Factor' scenario introducing extreme conditions to test while the user was sitting. For example, for face and iris recognition, we tested the device in a dark room with low lighting (around 4–5 lx approx); while for fingerprint recognition, we asked the user to dip their finger into a glass of water before attempting the authentication.

Table 6 presents the total False Reject Rates (FRR), the proportion of times a biometric system fails to grant access to an authorised person, found from each scenario. Here, the False Reject Rate is an outcome where the result was not a successful authentication, and this includes an unsuccessful recognition, user interaction errors, user cancelled or invalid sample capture.

This preliminary test begins to show the foundations of the model by testing several of the factors presented here including 'Modality', 'Scenarios', 'Environments', 'Users' and 'Hardware'. We argue that part of the increase FRR seen in the sitting scenario was due to this being the first scenario the user was presented and asked to authenticate themselves in and highlights the relationship between the users and hardware while the users adjust themselves to the current setup.

We also see how introducing a challenging condition (Factor) and therefore altering the environment can cause alternations in the performance, demonstrating a link between the modality and environments. We observe how darkening the lighting conditions saw a significant decrease of the FRR scores (Table 6) which we predict is due to the phone's use of an infrared camera which is able to focus more without disturbances and influences from any external light sources, highlighting the relationships between the modality, environment and scenario.

We investigate the acceptability asking users their preferred modality post-experiment. Only 13% of the participants confirmed a preference for iris, which likely reflects the FRR found while using the iris modality despite the scenario. For the other modalities under test, 65% of participants had a preference for fingerprint, while 20% for face and 2% for voice.

We begin to examine the 'mobility' relationship by introducing 'motion' scenarios and slightly against our expectations found the FRR tended to drop slightly in these 'motion' scenarios although this could be a consequence of the users' habitation with the device.

This initial practical demonstration shows that the factors we have identified are relevant at the beginning to assess the performance of biometrics on mobile systems and the remaining elements will become evident in future work.

13 Conclusion

In this paper, we have identified seven factors that are the core binning categories of performance alterations in mobile biometric systems. These factors are Users, Modality, Environments, Diversity of Scenarios, System Constraints, Hardware and Algorithms. As we have noted, these seven factors have significant overlap with one another. As an example, algorithms will require a particular hardware set-up to be able to function as expected. Also, the willingness of users to engage with a system is going to be profoundly affected by the modality and environment used. El-Abed *et al*. [35] stated that "evaluating biometric systems constitutes one of the main challenges in this research field". They also conclude by stating that "the main drawback of the widespread use of biometric technology is the lack of a generic evaluation methodology that evaluates biometric

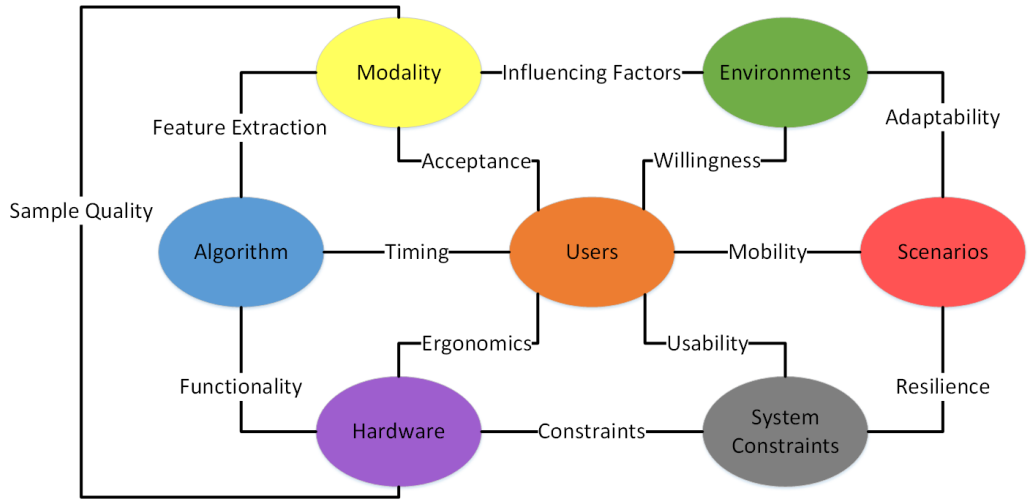

Fig. 4. Model Showing the Potential Relationships (Connections) between Factors

TABLE 4
Defining the Relationships Identified within the Model

Relationship	Definition [52], [53]	Measure
Influencing Factors	Any factor that affects the observed performance	Data Binning
Adaptability	The quality of being able to adjust to new conditions	Quantitative + Qualitative
Resilience	The capacity to recover quickly from difficulties	Qualitative
Constraints	A limitation or restriction	Data Binning
Functionality	The quality of being suited to serve a purpose well	Quantitative + Qualitative
Feature Extraction	The process of extracting information from data intending to be informative	Quantitative
Acceptance	Allowing a transaction using a specific modality	Qualitative
Willingness	The state of being prepared to operate within a particular environment	Qualitative
Mobility	The ability to move freely and easily	Quantitative + Qualitative
Usability	Measure of effectiveness, efficiency, and satisfaction	Quantitative + Qualitative
Ergonomics	The efficiency of the solution when being operated and handled by the user	Quantitative + Qualitative
Timing	The time that is taken by a process, activity, or a person doing it	Quantitative
Sample Quality	The fitness of a biometric sample to accomplish the comparison decision	Quantitative

systems taking into account: performance, users' acceptance and satisfaction, data quality and security aspects".

Each of the defined factors could easily extend and expand to incorporate more detail. However, the aim here is to highlight the central concepts that create a foundation acting as a *parent node* if we were to use tree terminology. For example, we have identified *Users* as one of the factors, but many subsections will occur out of this, including interaction and acceptance both of these are significant areas that could arguably be a factor of their own, but we still identify *Users* as the central area incorporating these.

It is interesting to note that some of the factors we have identified fit directly into the HBSI model proposed by Elliott *et al*. [21]. Here, Human, Sensor and Biometric System become *Users*, *Hardware* and *System Constraints* respectively. As we are trying to identify factors that affect performance in different ways, it is reassuring to note that HBSI model is still present and preserved within this updated model which goes beyond usability aspects to define performance.

Comparing the factors presented here with the ones originally provided by Mansfield *et al*. [2] and included within the ISO/IEC 19795-1:2006 [3], we can see that there exists a similarity between them.

- Population Demographics ↔ *Users*
- Application ↔ *System Constraints*
- User Physiology ↔ *Users*
- User Behaviour ↔ *Users*
- User Appearance ↔ *Users*
- Environmental Influences ↔ *Environments*
- Sensor and Hardware ↔ *Hardware*
- User Interface ↔ *System Constraints*

The factors from Mansfield *et al*. only covers four of the seven factors we have presented. It confirms the idea of *Users* being one of the most influential factors and demonstrates how the *Users* consideration can extend into subsections incorporating what Mansfield *et al*. have previously identified. The factors we have added are *Modality*, *Diversity of Scenarios* and *Algorithms* and together these make the core factors for mobile biometric performance.

Looking back at Jain *et al*. [1] seven factors for assessing biometric traits we mentioned that we are going to be taking influence from them in terms of *performance*, *acceptability* and *circumvention*. We have incorporated *acceptability* within the

TABLE 5
Examples of Suggested Methods for Collecting Model Relationship Data

Relationship	Collection Suggestion	Explanation	Examples
Influencing Factors	Literature	Explore current influencing factors	[Illumination, Noise, Wearing Glasses]
Adaptability	Algorithmic	Measure standard performance rates[1] in various environments and scenarios	While in "Environment 1" FAR increased to 9%
Resilience	Algorithmic	Measure standard performance rates[1] across a range of challenging conditions	In a challenging condition FRR was 34%
Constraints	Literature	Explore current hardware that can be supported and usable	[Identification, Off-Device Processing, 2 Seconds to Authenticate]
Functionality	Statistical	Explored with analysis of using different algorithms and hardware	"Algorithm 1" achieved 88% successful matches and "Algorithm 2" achieved 92%
Feature Extraction	Algorithmic	Measure of how well algorithm performs at extracting features	Extraction was able to find a total of 8 total features
Acceptance	Questionnaire	Survey of users	Survey revealed that 80% of users would allow a transaction to happen with chosen modality
Willingness	Questionnaire	Survey of users	Only 20% of surveyed users would be happy to use this verification method in the chosen environment
Mobility	Statistical	Explored with analysis of performance in motion scenarios	While in motion the FRR was 13%
Usability	Questionnaire + Interaction	Survey of users and interaction measures	74% of users were satisfied and it took 10 seconds to read each of the task prompts
Ergonomics	Questionnaire + Interaction	Survey of users and interaction measures	67% of users were comfortable and managed to complete the task within 35 seconds
Timing	Experimental	Device in operation should be capable of capturing timings	Authentication took an average of 7 seconds to complete
Sample Quality	Algorithmic	Sample quality can be measured to ISO and similar standards for some modalities	Sample quality score achieved was 81

[1] Standard Performance Rates = FRR, FAR, FTA, FTE

TABLE 6
False Reject Rate of modalities on the Samsung Galaxy S9 in a variety of scenarios

	Fingerprint	Face	Iris
Sitting	26%	8%	28%
Standing	6%	10%	17%
Treadmill	4%	9%	27%
Corridor	7%	7%	29%
Factor	Wet - 77%	Dark - 1%	Dark - 25%

Users factor and *circumvention* within the *System Constraints* factor. The definition of *performance* that they provided talks about the "accuracy, speed, and robustness" and these are fundamental concepts. However, we believe that this needs updating to incorporate the other factors presented here to provide more assurance for a mobile context.

In defining the factors, we provide an informative overview to developers, implementers and testers of biometrics systems, enabling the binning of performance alterations within one of these factors. We expect categorical overlaps, so it is quite likely that a performance alteration will have many factors contributing to the observed effect.

14 Future Work

Having identified the seven factors; the next step is to find ways to mitigate these effects suitably. Within a mobile context, it is impossible to test every possible usage outcome. Therefore, we need to develop an approach that provides us with the confidence that the system is *fit for purpose*. Adopting this change will likely mean a modification to the testing strategy that the community is currently familiar with as it will require detailed testing that includes more situations now available in a mobile context. Other metrics for performance will need adding into standard testing procedures, including ways of measuring usability from the HBSI model and data quality. These changes will begin to bring more confidence into results from biometrics studies and allow users to feel more comfortable while interacting with a biometric system. More research will need to be conducted to identify the quality of biometric samples under various conditions with more significant influence given to the collection environment.

Research is shifting to accommodate this change from a fixed to a more mobile system and exploring new opportunities and situations for mobile biometrics. Hopefully, the identified factors will help to pave the way for future research to focus on some of these critical areas and allow for future biometric systems to have a high level of performance that provides the *fit for purpose* assurance. This work forms the first steps in trying to design a suitable framework that can assess performance. The next step is to find ways of

turning these factors into measurable metrics that can be used to help both analyse, compare and visualise results more effectively.

REFERENCES

[1] A. K. Jain, R. Bolle, and S. Pankanti, *Biometrics: personal identification in networked society*. Springer Science & Business Media, 2006, vol. 479.

[2] A. J. Mansfield and J. L. Wayman, "Best practices in testing and reporting performance of biometric devices," *Centre for Mathematics and Scientific Computing, National Physical Laboratory Teddington, Middlesex, UK*, 2002.

[3] ISO/IEC 19795-1:2006, "Information technology – biometric performance testing and reporting – part 1: Principles and framework," International Organization for Standardization, Standard, Apr. 2006.

[4] K. Hollingsworth, K. W. Bowyer, and P. J. Flynn, "Pupil dilation degrades iris biometric performance," *Computer vision and image understanding*, vol. 113, no. 1, pp. 150–157, 2009.

[5] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 721–732, 1997.

[6] R. Cappelli, M. Ferrara, and D. Maltoni, "The quality of fingerprint scanners and its impact on the accuracy of fingerprint recognition algorithms," in *International Workshop on Multimedia Content Representation, Classification and Security*. Springer, 2006, pp. 10–16.

[7] Oxford Dictionary, "factor — definition of principle in english by oxford dictionaries," 2018, accessed 2018-07-16. [Online]. Available: https://en.oxforddictionaries.com/definition/factor

[8] D. Lee, "San francisco is first us city to ban facial recognition," May 2019, accessed 2019-07-10. [Online]. Available: https://www.bbc.co.uk/news/technology-48276660

[9] A. Buriro, Z. Akhtar, B. Crispo, and S. Gupta, "Mobile biometrics: Towards a comprehensive evaluation methodology," in *Security Technology (ICCST), 2017 International Carnahan Conference on*. IEEE, 2017, pp. 1–6.

[10] R. Blanco-Gonzalo, R. Sanchez-Reillo, O. Miguel-Hurtado, and J. Liu-Jimenez, "Usability analysis of dynamic signature verification in mobile environments," in *Biometrics Special Interest Group (BIOSIG), 2013 International Conference of the*. IEEE, 2013, pp. 1–9.

[11] Z. Syed, J. Helmick, S. Banerjee, and B. Cukic, "Effect of user posture and device size on the performance of touch-based authentication systems," in *High Assurance Systems Engineering (HASE), 2015 IEEE 16th International Symposium on*. IEEE, 2015, pp. 10–17.

[12] B. Fernandez-Saavedra, R. Sanchez-Reillo, R. Ros-Gomez, and J. Liu-Jimenez, "Small fingerprint scanners used in mobile devices: the impact on biometric performance," *IET Biometrics*, vol. 5, no. 1, pp. 28–36, 2016.

[13] F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "Quality measures in biometric systems," *IEEE Security & Privacy*, vol. 10, no. 6, pp. 52–62, 2011.

[14] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern recognition*, vol. 38, no. 12, pp. 2270–2285, 2005.

[15] M. He, S.-J. Horng, P. Fan, R.-S. Run, R.-J. Chen, J.-L. Lai, M. K. Khan, and K. O. Sentosa, "Performance evaluation of score level fusion in multimodal biometric systems," *Pattern Recognition*, vol. 43, no. 5, pp. 1789–1800, 2010.

[16] D. Gafurov, "A survey of biometric gait recognition: Approaches, security and challenges," in *Annual Norwegian computer science conference*, 2007, pp. 19–21.

[17] K. Ito and T. Aoki, "Recent advances in biometric recognition," *ITE Transactions on Media Technology and Applications*, vol. 6, no. 1, pp. 64–80, 2018.

[18] P. Yan and K. W. Bowyer, "Biometric recognition using 3d ear shape," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 8, pp. 1297–1308, 2007.

[19] R. Blanco-Gonzalo, R. Sanchez-Reillo, L. Martinez-Normand, B. Fernandez-Saavedra, and J. Liu-Jimenez, "Accessible mobile biometrics for elderly," in *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. ACM, 2015, pp. 419–420.

[20] R. Sanchez-Reillo, R. Blanco-Gonzalo, J. Liu-Jimenez, M. Lopez, and E. Canto, "Universal access through biometrics in mobile scenarios," in *2013 47th International Carnahan Conference on Security Technology (ICCST)*. IEEE, 2013, pp. 1–6.

[21] S. Elliott, E. Kukula, and N. Sickler, "The challenges of the environment and the human/biometric device interaction on biometric system performance," in *International Workshop on Biometric Technologies-Special forum on Modeling and Simulation in Biometric Technology, Calgary, Alberta, Canada*, 2004.

[22] C. Lunerti, R. M. Guest, R. Blanco-Gonzalo, R. Sanchez-Reillo, and J. Baker, "Environmental effects on face recognition in smartphones," in *Security Technology (ICCST), 2017 International Carnahan Conference on*. IEEE, 2017, pp. 1–6.

[23] Y. Gong, "Speech recognition in noisy environments: A survey," *Speech communication*, vol. 16, no. 3, pp. 261–291, 1995.

[24] T. Yamada, M. Kumakura, and N. Kitawaki, "Performance estimation of speech recognition system under noise conditions using objective quality measures and artificial voice," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, 2006.

[25] R. Guest, M. Brockly, S. Elliott, and J. Scott, "An assessment of the usability of biometric signature systems using the human-biometric sensor interaction model," *International Journal of Computer Applications in Technology*, vol. 53, no. 4, pp. 336–347, 2016.

[26] S. Elliott, M. Mershon, V. Chandrasekaran, and S. Gupta, "The evolution of the hbsi model with the convergence of performance methodologies," in *2011 Carnahan Conference on Security Technology*. IEEE, 2011, pp. 1–4.

[27] M. Brockly, R. Guest, S. Elliott, and J. Scott, "Dynamic signature verification and the human biometric sensor interaction model," in *Security Technology (ICCST), 2011 IEEE International Carnahan Conference on*. IEEE, 2011, pp. 1–6.

[28] R. Bhagavatula, B. Ur, K. Iacovino, S. M. Kywe, L. F. Cranor, and M. Savvides, "Biometric authentication on iphone and android: Usability, perceptions, and influences on adoption," *Proceedings 2015 Workshop on Usable Security*, 2015.

[29] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani, "Hmog: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 877–892, 2016.

[30] K. Nguyen, C. Fookes, R. Jillela, S. Sridharan, and A. Ross, "Long range iris recognition: A survey," *Pattern Recognition*, vol. 72, pp. 123–143, 2017.

[31] G. Doddington, W. Liggett, A. Martin, M. Przybocki, and D. Reynolds, "Sheep, goats, lambs and wolves: A statistical analysis of speaker performance in the nist 1998 speaker recognition evaluation," National Institute of Standards and Technology Gaithersburg MD, Tech. Rep., 1998.

[32] N. Yager and T. Dunstone, "Worms, chameleons, phantoms and doves: New additions to the biometric menagerie," in *2007 IEEE Workshop on Automatic Identification Advanced Technologies*. IEEE, 2007, pp. 1–6.

[33] N. Popescu-Bodorin, V. E. Balas, and I. M. Motoc, "The biometric menagerie–a fuzzy and inconsistent concept," in *Soft Computing Applications*. Springer, 2013, pp. 27–43.

[34] R. Blanco-Gonzalo, N. Poh, R. Wong, and R. Sanchez-Reillo, "Time evolution of face recognition in accessible scenarios," *Human-centric Computing and Information Sciences*, vol. 5, no. 1, p. 24, 2015.

[35] M. El-Abed, R. Giot, B. Hemery, and C. Rosenberger, "A study of users' acceptance and satisfaction of biometric systems," in *Security Technology (ICCST), 2010 IEEE International Carnahan Conference on*. ieee, 2010, pp. 170–178.

[36] A. Wojciechowska, M. Choraś, and R. Kozik, "The overview of trends and challenges in mobile biometrics," *Journal of Applied Mathematics and Computational Mechanics*, vol. 16, no. 2, 2017.

[37] C. Riley, K. Buckner, G. Johnson, and D. Benyon, "Culture & biometrics: regional differences in the perception of biometric authentication technologies," *AI & society*, vol. 24, no. 3, pp. 295–306, 2009.

[38] X. Qu, D. Zhang, G. Lu, and Z. Guo, "Door knob hand recognition system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 11, pp. 2870–2881, 2017.

[39] O. Miguel-Hurtado, R. Blanco-Gonzalo, R. Guest, and C. Lunerti, "Interaction evaluation of a mobile voice authentication system," in *Security Technology (ICCST), 2016 IEEE International Carnahan Conference on*. IEEE, 2016, pp. 1–8.

[40] M. A. Sasse, "Red-eye blink, bendy shuffle, and the yuck factor: A user experience of biometric airport systems," *IEEE Security & Privacy*, vol. 5, no. 3, 2007.
[41] T. Feng, Z. Liu, K.-A. Kwon, W. Shi, B. Carbunar, Y. Jiang, and N. Nguyen, "Continuous mobile authentication using touchscreen gestures," in *Homeland Security (HST), 2012 IEEE Conference on Technologies for*. IEEE, 2012, pp. 451–456.
[42] D. Buschek, A. De Luca, and F. Alt, "Improving accuracy, applicability and usability of keystroke biometrics on mobile touchscreen devices," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 1393–1402.
[43] L. Jain, J. V. Monaco, M. J. Coakley, and C. C. Tappert, "Passcode keystroke biometric performance on smartphone touchscreens is superior to that on hardware keyboards," *International Journal of Research in Computer Applications & Information Technology*, vol. 2, no. 4, pp. 29–33, 2014.
[44] J. Chakrabarthy, "Fingerprint scanner on phones: History & evolution, but do we really need that?" 2016, accessed 2018-05-29. [Online]. Available: https://www.igadgetsworld.com/fingerprint-scanner-history-evolution-but-do-we-really-need-that/
[45] Apple, "About touch id advanced security technology," 2017, accessed 2018-05-29. [Online]. Available: https://support.apple.com/en-gb/HT204587
[46] D. D. Hwang and I. Verbauwhede, "Design of portable biometric authenticators-energy, performance, and security tradeoffs," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 4, pp. 1222–1231, 2004.
[47] R. R. Schaller, "Moore's law: past, present and future," *IEEE spectrum*, vol. 34, no. 6, pp. 52–59, 1997.
[48] E. Cantó-Navarro, M. López-García, and R. Ramos-Lara, "Floating-point accelerator for biometric recognition on fpga embedded systems," *Journal of Parallel and Distributed Computing*, vol. 112, pp. 20–34, 2018.
[49] P. Hu, H. Ning, T. Qiu, Y. Xu, X. Luo, and A. K. Sangaiah, "A unified face identification and resolution scheme using cloud computing in internet of things," *Future Generation Computer Systems*, vol. 81, pp. 582–592, 2018.
[50] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1701–1708.
[51] N. Ortiz, R. D. Hernandez, R. Jimenez, M. Mauledeoux, and O. Aviles, "Survey of biometric pattern recognition via machine learning techniques," *Contemporary Engineering Sciences*, vol. 11, no. 34, pp. 1677–1694, 2018. [Online]. Available: https://doi.org/10.12988/ces.2018.84166
[52] English Oxford Living Dictionaries, "English dictionary, thesaurus, & grammar help — oxford dictionaries," 2018, accessed 2018-07-19. [Online]. Available: https://en.oxforddictionaries.com/
[53] ISO/IEC 2382-37:2017, "Information technology — vocabulary — part 37: Biometrics," International Organization for Standardization, Standard, 2017.
[54] ISO/IEC 29794-1:2016, "Information technology – biometric sample quality – part 1: Framework," International Organization for Standardization, Standard, Jan. 2016.

Matthew Boakes obtained his BSc degree in Computer Science from the University of Kent, Canterbury, the United Kingdom in 2017. He is currently working on his PhD at the University of Kent, School of Engineering and Digital Arts on a project titled 'A Performance Assessment Framework for Mobile Biometrics'. He is part of the Intelligent Interactions research group and a member of the Kent Interdisciplinary Research Centre in Cyber Security (KirCCS). His research interests include biometrics, artificial intelligence and pattern recognition. He is currently a BSI Biometrics (IST/44) committee member.

Richard Guest received his Ph.D. degree in Electronic Engineering from the University of Kent, Canterbury, U.K. in 2000. Since then, he has been a member of academic staff (currently Reader in Biometric Systems Engineering) with the University of Kent, where he is also the Deputy Head of the School of Engineering and Digital Arts. His research interests include image processing and pattern recognition, specialising in biometric and forensic systems, particularly in the areas of behavioural and image information analysis, standardisation, and usability.

Farzin Deravi received the B.A. degree in Engineering Science and Economics from the University of Oxford, U.K., in 1981, an M.Sc. degree in Communications Engineering from Imperial College London, U.K., in 1982, and a Ph.D. degree in Electronic Engineering from the University of Wales, Cardiff, U.K., in 1988. He is currently Professor of Information Engineering and Head of School of Engineering and Digital Arts at the University of Kent. His research interests include the fields of pattern recognition and signal processing and their application in security and healthcare.

Barbara Corsetti received a Bachelor Degree in Clinical Engineering in 2014 and two years later she obtained a Master degree in Biomedical Engineering. During her studies, she got experience in biomedical signals, biomechanics and biofluid dynamics. On September 2017 she joined the University Group for Identification Technology (GUTI) at University Carlos III of Madrid, where she is currently doing her PhD within the AMBER project on the evaluation of biometrics systems.