International Journal of Information Security and Privacy
Volume 17 • Issue 1

Trust and Voice Biometrics Authentication for Internet of Things

Alec Wells, York St. John University, UK
Aminu Bello Usman, York St. John University, UK*

ABSTRACT

In recent years, IoT adoption has been higher, and this causes lots of security concerns. One of the fundamental security concerns in IoT adoption is the question, "Are you who you say you are?" Thus, authentication forms the gateway for a secure communication system with IoT. So far, the human voice is one of the most natural, non-intrusive, and convenient behavioural biometric factors compared to other biometric authentication methods. Despite the non-intrusive characteristics of voice as a biometric authentication factor when accessing IoT technologies, there is a concern of a general societal trust and distrust with IoT technology and the risk of theft of users' data and imitation. This study derived a realistic trust evaluation model that incorporates privacy, reliability, security, usability, safety, and availability factors into a trust vector for a flexible measurement of trust in the user accessing IoT technologies.

KEYWORDS

Biometric Authentication, Internet of Things, Trust, Voice Biometric Authentication

1. INTRODUCTION

The Internet-of-Things (IoT) brought in the rapid expansion of wearable and connected devices that offer new ways to develop smart applications that drive efficiencies, engage users, and develop new businesses with more significant insights at the Intelligent Edge and on the cloud. These technologies are invading every aspect of our lives, including our homes, offices, and health, and the military and aviation, etc. Increasing in elderly population IoT technologies are being deployed to improve the quality of life and make life easier for seniors. Examples of IoT applications include using voice command to turn on a light or using a phone to turn on or off light or using smart watch to help track sleep pattern or fitness device to transmit data to a doctor or a nurse so they can help monitor your activity. Other examples include using a smart speaker like Amazon Echo or Google Home to play music, hear the weather forecast, or even get help with a recipe.

Having access to IoT devices requires authentication for secure communication. Thus, authentication forms the gateway for a secure communication system with IoT since it is the most common method that allows the users to have access to IoT devices.

The most common authentication methods used in IoT include: the use of a piece of information the user knows (e.g., secret questions, passwords, PINs), owns (e.g., ID cards, mobile phones, tokens)

DOI: 10.4018/IJISP.322102 *Corresponding Author

or information they inherited (e.g., biometrics, fingerprints, iris scans, signatures) (Fu, 2015). Strong authentication systems help to reduce potential fraudsters and other hackers from gaining access to sensitive information on IoT technologies. Biometrics presents an intriguing window of opportunity to improve IoT usability and security and can play a critical role in protecting a wide range of developing IoT devices to address security challenges and are predicted to target future applications of IoT (Yang et al., 2021).

The human voice is one of the most natural, non-intrusive and convenient behavioural biometric factors in comparison to other biometric factors. Subsequently, the usage of voice biometric authentication in IoT technology is being heavily considered as a promising means of IoT authentication for many reasons (Ortega-Garcia et al., 2004). These include not requiring the user to remember any pins or passwords, users constantly being verified - hence fraudsters are more easily caught, and verification being done over standard telephones lines through already implemented infrastructure (Tupman, 2018). Also, unlike knowledge-based factors, it does not fall to the user to create a strong authentication, as shown with passwords, where many users use weak or even default passwords allowing hackers easy access to accounts whether it be children from lack of cognitive ability or adults who don't understand the importance of strong cyber security (Choong et al., 2019).

With voice biometric authentication being such a promising development when it comes to secure user authentication, it is of interest to consider the user's current perspectives of the authentication to better understand if users would first be willing to use the technology especially when compared with traditional means of authentication. As despite all the advantages and non-intrusive characteristics of voice biometric over other biometric features, voice biometric authentication has been brought under scrutiny for many reasons including: the accuracy of biometric data, the risk of theft and imitation and a general societal feeling of distrust with technology to handle their privacy-sensitive information and biometric data securely, preventing full adoption of voice biometric authentication systems. Other lingering concerns about the security of voice biometric systems include potential breaches of privacy, adversary attacks such as Spoofing attack (Marcel et al., 2019), presentation attacks (Korshunov & Marcel, 2017) and Replay-attacks (Lavrentyeva et al., 2017) may cause a user distrust to technology.

As it currently stands, for new technologies such as voice biometric technology to be adopted; users must overcome the challenges of trusting said technology. Thus, trusting technology is a critical factor for the successful adoption of a new technology, products, or service (Hoffman et al., 2006). Other factors such as lack of clarity, confidence, poor user experience and expectation of the technology, may prompt concerns about whether to adopt the new technology of which will increase or decrease along with users trust of the technology. Subsequently, it will be important to understand which method of user-based authentication mechanism could facilitate trust establishment between user and technology from the user's perspective?

Although, when it comes to trusting technology, users' perceptions change over the course of time through continued use of technology allowing perceptions and opinions to change. However, since trusting technology beyond their functionality and capacity can present a high risk, cost, and compromise to user privacy and personal security, it would be interesting to understand which trust evaluation model can be employed for flexible measurement of trust (in the context of availability, security, usability, privacy, reliability, willingness to use and security) between the user and security-based authentication mechanism to access technology?

Our contribution in the work is three-fold:

1. Motivated by the expanded trust model proposed by [9, pp. 95–101], we derived a realistic trust evaluation model that incorporates privacy, reliability, security, usability, safety, and availability factors into a trust vector, for a flexible measurement of trust in the context of user accessing the technology.
2. Based on the derived trust model we experiment using mixed method whether the users are willing to trust voice biometric authentication method over PIN, fingerprint and token-based

authentication and hence would be inclined to adopt and utilize it as a means of user authentication to access technology.
3. We applied Kruskal-Wallis H test and the post-hoc test to understand which authentication method the user trusts, based on statistical significance and which groups were found to have that statistical difference.

Understanding users' trust opinions is crucial for building technology that meets users' needs and expectations, inspires trust, and drives adoption and long-term usage. Thus, the novelty of this work is cantered around the derivation of trust model that helps measure the level of user's trust on different authentications methods. For example, by understanding users' trust opinions, developers can design technology that correlates with users' expectations and preferences, thereby enhancing the user experience. This can result in an improved user experience, greater user satisfaction, and increased adoption rates. Further, when users have trust in a technology, they are more likely to adopt and routinely use it. Developers can design technology that inspires trust and confidence in users by gaining an understanding of users' trust perceptions using our proposed trust model.

The paper is structured as follows: Section two gives the discussion of related work around different authentication methods and related studies. Section 3 discusses the different elements of the trust model used and research design. Section 4 is a collection of the results of the study as well as a discussion. Section 5 is a conclusion and proposition of potential future work.

2. RELATED WORK

While the technology for voice biometric authentication has been around for years now, only in recent years has it seen huge developments primarily in IoT technologies. There are two main variations of voice biometric authentication: text-dependent and text-independent. Text-dependent systems require the same specific phrase to by said by the user they used to set up the voice print, often called a passphrase. The more common, text-Independent systems in contrast do not require the use of passphrases and instead the identification is often done without the user's knowledge (Microsoft, 2006).

Voice biometric authentication will require the user to give a sample of their speech via talking into a microphone. This speech is then converted into a voiceprint that is stored in a database, sometimes as a waveform, it can then be referred to and compared when a user wishes to gain entry and authenticate themselves (Krawczyk & Jain, 2005). The user will once again speak and have their voice recorded via a microphone or telephone, their speech is then converted into another voiceprint, which is compared against the one they provided prior. If the two voiceprints match, the user is then authenticated.

Voice biometrics has seen some applications in important services already. One such usage of voice recognition software is within banking authentication, which has been deployed by banks such as Barclays, Santander, among others. Barclays use voice recognition as a means to authenticate the identity of the customer calling their call centres. This is done by comparing the voice that calls them over the phone to the voice they have on file and identifying if there is a match, which has reduced average call time by 15% (Nuance no date). By 2017 the number of people enrolled was already 160 million, which is expected to grow to about 600 million by 2020 (Jones, 2018). This system, which is utilised by Barclays has been developed by the company Nuance, who have had their technology utilised by a wide variety of companies.

2.1. Knowledge-Based Factors

Knowledge-based factoids are based on information only the user should know, such as a username and password or a personal identification number (PIN). The two most widely used methods of users' authentication using KBA are: static (shared secrets) and instant (also known as dynamic KBA). Static KBA is based on a pre-defined or agreed set of questions or alternatively a shared

piece of secret information between the authentication parties involved. Mostly, static KBA factoids include questions such as what is your mother's maiden name? Or what is your date of birth etc, and is commonly used by email providers, banks, financial services or companies to authenticate users. On the contrary, instant KBA uses methods and algorithms to dynamically develop set of personal questions and answers to authenticate a user, meaning it does not require the user to have provided the questions and answers beforehand (Fu 2015). These dynamic questions provide randomized right and wrong answer choices based on data found for the subject by the KBA system. Regardless, in practical usage, both versions of KBA usually require a form of initial registration against an existing database to create the credentials. KBA then usually requires some online or remote access to the server to verify the factoids/credentials in the login mechanism (Chokhani, 2004).

2.2. Ownership-Based Factors

Ownership-based authentication factors are based on something the user has, such as cards, smartphones, or other tokens. For instance, one of the most prevalent examples of ownership-based factors are payment cards, utilised by banks that each possess a unique combination of numbers and security information from one another. Another example of ownership-based factors is the usage of tokens which are issued to the user when they need to sign in.

Payment cards are an extremely common form of ownership-based factors and are usually issued by banks. A bank card has multiple strings of data, such as a unique string of numbers, an expiry date and a security code that is tied to a user's band account. Bank cards can come in many different forms, with the most common being credit and debit cards. Similarly, many banks also use tokens/ one-time use passwords to authenticate users and the server. Authentication apps and messages are being used for a variety of online accounts to be used in conjunction with passwords as a form of two-factor authentication, some examples include the google authenticator and windows authenticator apps. Alternatively, mobile phones can be used as tokens via Bluetooth wireless communication, using the phone token as a challenge-response protocol (Kunyu et al., 2009).

Another form of ownership-based factors is smart cards, which are also known as an integrated circuit card (ICC card) - an electronic authorization device, used to control access to a resource. The ICC card is usually a credit card-sized card with an embedded integrated circuit (IC) chip (ISO/IEC, 2007). A smart card can either be in a form of card with a metal contact to electrically connect either via an internal chip, via contactless means or via both forms (Kuo & Lo, 1999). Smart card contains users' authentication, small data storage, and application processing components to perform Input/output (I/O) functions. In terms of applications, most organisations used smart cards for single sign-on (SSO) for pass-through authentication system. Other forms of ownership-based factors include NFC (Chen et al., 2010), RFID (Lim & Kwon, 2006), Hardware-token (Shablygin et al., 2013), and cell-phone.

2.3. Inheritance-Based Factors

The human body provides indispensable sources of distinct features, which are suitable for the function of authentication systems or recognition. The use of such distinctive features of human body or person's inheritance characteristics is referred to as Inheritance-based authentication process which employed different modalities and factors such as fingerprints, the iris, voice recognition and the face. There are different algorithms and techniques to extract the physical and characteristics for inheritance or biometric traits such as hand geometry, palmprints, the ears, the mouth, or the nose. In the current state of the art, the analysis of the retinal vascular pattern with respect to individuals (pattern of blood vessels), appears to be one of the main sources of biometric features in methods such as retina scan or vein matching (Rigas et al., 2016). Multiple biometric features can even be combined using biometric fusion to improve security and has been shown to improve further using deep learning to outperform traditional state-of-art-methods (Arora & Bhatia, 2021). Of biometric fusion, it is found

that across 200 papers, score-level fusion is the most general technique, being done with a newly generated image is matched against a previous image to create a score-level fusion (Bala et al., 2021).

Other forms of inheritance-based traits that enfold behavioural distinctive characteristics and are partially connected with brain activity include the keystroke dynamics, voice recognition, speech analysis and eye movement driven biometrics. Unlike knowledge or ownership factors, that are about what the user knows or has, inherence-based factors are based on the user themselves – using the body for measurements and characteristics. inheritance-based authentication can be split into two different categories. The first is physical biometrics, that uses physiological features of the human body for users' authentication, this includes methods such as using a fingerprint or iris scanning. Alternatively, there are behavioural biometrics, which utilise a pattern of behaviour that is specific to the user, this includes methods such as voice recognition or could be the rhythm they usually type on a keyboard (Chrobok Mateusz, 2020).

3. TRUST MODEL, METRICS, AND RESEARCH FRAMEWORK

Trust in technology shares many similar principles with trust in a social context, however, the two perceptions of trust are not identical. For instance, whereas with trust in a social context, a trustee relies on an individual to behave in a reliable manner, though an individual can act on their own and has free will hence be unpredictable, unlike a technology which should consistently perform tasks in a predictable manner, it may have issues with reliability suffering failures (McKnight, 2005). In addition, a user can often expect a technology to fulfil certain functionalities and produce correct results, whereas in a social context when a user depends on another human, they can often only expect that task to be done to the best of that person's capability (Barber, 1983). Furthermore, when trusting in a social context with people, a user has to take in that person's desire to help and availability to help, unlike with a technology in which that technology is available at all times (Rempel et al., 1985).

Trust is commonly defined as a confident expectation about a situation leading to willingness to accept vulnerabilities that arise from risk and situational uncertainty (Dietz, 2011). Trusting technology beyond their functionality and capacity can present high risk, cost as well as compromise to user privacy and personal security.

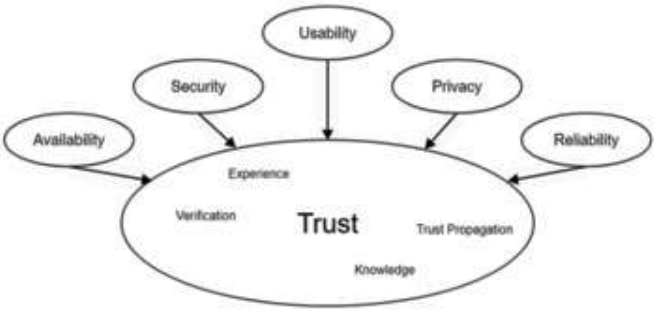
In the first place, trust has to do with the belief, uncertainty, intention, and willingness to trust or not to trust (Usman & Gutierrez, 2019). These attributes are behavioural characteristics which cannot be accurately measured and predicted with a high degree of accuracy. Measuring trust level of a user in a social setting or an agent in a distributed system can be a complex process because of the dynamic and unpredictable nature of trust. Quantitative metrics (Cruz et al., 2019), qualitative metrics (Patent & Searle, 2019), fuzzy metrics, or an amalgamation are used to measure trust levels with the trust model and metrics has its own specific characteristics and requirements; nonetheless, their pattern and abstract scheme can be generalized. Hence, based on our research behind what defines trust, we used the trust model proposed by (Hoffman et al., 2006) over other similar models given it is a reasonably well cited article and expands on many metrics laid out by older trust models as shown in Figure 1.

We define trust as the expectation and experience that a technology will provide the user with the sense of security, reliability, and confidence. With this definition, to derive users' trust, we identified five main key components of the trust definition. The five keywords are: usability, availability, security, privacy, and reliability. For a realistic adoption of the Hoffman, et al trust model, we identified four other internal metrics that can contribute to the derivation of our trust model. The metrics are user experience, recommendation (trust propagation), knowledge and verification.

3.1. Trust Metrics and Elements

We described the presented trust metrics and elements from Figure 1:

Figure 1. Expanded trust model (Hoffman et al., 2006)



- **Availability:** How widespread and utilised the technology is. From a user's point of view, availability means that they can access it whenever they require, regardless of specialist equipment being available to the authenticate user when needed, being widespread and being a common example of other authentication techniques.
- **Security:** How secure the technology is from potential attacks. From a user's point of view, security is important in trusting that the voice biometric authentication will perform the users' intended functions with relative security, that the authentication method is not easily hacked, tampered with and that the method is able to differentiate the user with a degree of accuracy.
- **Usability:** How easy a user can utilise the features of the authentication. From a user's point of view usability means they can use it without confusion meaning it is easy to use, easy to learn and is accessible to different needs such as disability.
- **Privacy:** How the method keeps user's sensitive information secure and protect the anonymity of a users' identity. From a user's perspective this means protecting their privacy from others, preventing others from seeing the contents of their data and allows the user to remain anonymous.
- **Reliability:** The reliability of the technology in the eyes of the user as how the technology can consistently perform and function as expected by the user, performed the same each time it was used and will continue to perform as expected in further uses.

As can be seen, the trust model presented in Figure 1, incorporates aspects privacy, reliability, security, usability, safety, and availability mechanisms, as well as user privacy concerns, user experience, and user knowledge about the technology. The model also identified four key elements that contributed to formulation of trust.

The experience users have had with the method, such as have they used it many times before, do they use similar methods that are based on the same factors, or do they use different authentication methods often.

The verification the method provides the user with feedback provided by the method to show it has been processed the authentication request correctly, an error had occurred or that it had been set-up correctly.

The knowledge of the method a user has, such as the understanding they have of the authentication method itself, how the overall authentication process works and why it is used.

The propagation of the method, or what experiences about the method they have shared, or been shared, such as good experiences, bad experiences or just their general perception of the method reputation.

An example of how these factors interact could be from the user continuing to use the technology and that technology proving to be reliable and produce the same consistent results each time, can give the user a good experience of the technology and build a more trusting relationship. Alternatively, a

user might hear from a colleague about how easy a technology is to use, and that trust propagation helps to build a foundation level of trust with the user and that technology. The aspects of trust from the model can be used to inform technology developers of the trust issues users will have with a system and in turn when developing new systems can implement systems based on the aspects of trust to improve the levels of trust the user has with the system.

3.2. Research Design

In total, 60 participants took part in the study who were aged between 18-60 of which, 46 were male and 14 were female. The participants were gathered around the university campus and asked if they could assist in the study, after being verbally briefed about the content, they filled out a consent form to opt into the study, which would take between 5 and 10 minutes. All participants were English speakers and had varying previous amounts of exposure to the chosen authentication methods. The study was conducted over the period of a month from the end of September 2020 to the end of October 2020. A between-subjects study design was utilised hence, each participant only used one of the authentication methods, meaning there were 15 participants for each authentication method.

To test the four-authentication covered by this study (PIN, fingerprint, token and voice biometrics) an android Samsung Galaxy S9 phone was used alongside a standard university computer running windows 10 operating system. To test PIN & fingerprint, participants used the built-in settings of the S9 to set up their authentication and then unlock the phone. Control groups testing voice biometric authentication, used google assistant to unlock the phone via voice commands after setting up a voice profile. Meanwhile participants utilising tokens, were provided a token through a windows application that they first set up and then used that token to unlock the phone.

After utilising the authentication method, a proxy measure of trust was gathered via the questionnaire. The questionnaire's questions were developed from the 'Generic Trust Model' using the core aspects of availability, security, usability, privacy and reliability which each contribute to the acquisition of trust. The question also was developed from the connections identified, which also affect the level of trust a user would have: the experience of the technology, the verification/feedback provided by the technology, the knowledge/understanding of how the technology functions and the propagation of good or bad experiences with the technology. Participants answered 3 questions for each of the 9 categories, ranking their opinions on a scale of 1-5, from strongly disagree to strongly agree.

4. RESULTS DISCUSSION

We analysed the results using the Kruskal-Wallis H test, to find which results were deemed statistically significant and test the hypothesis:

Users have varying degree of trust about user-based biometric authentication method to access technology based on the chosen trust evaluation model.

Users may be found to be willing to trust technology and voice-based biometrics as a method of user authentication.

or the null hypothesis:

Users have the same degree of trust about user-based biometric authentication method to access technology based on the chosen trust evaluation model.

Users may be found to not be willing to trust technology and voice-based biometrics as a method of user authentication.

We used Kruskal-Wallis test instead of one-way ANOVA as there is no assumption that our data would have a normal distribution, hence we ran the non-parametric Kruskal-Wallis H test. To determine which specific groups had a statistical significance to one another, we then ran a post-Hoc test on the groups that had a statistical significance. We applied the mean rank as the average of the ranks for all observations within each sample of the collected data. Since we are using the Kruskal-Wallis H test (McKight & Najab, 2010), we used SPSS to rank the combined samples by assigning the smallest observation a rank of 1, the second smallest observation a rank of 2, and so on. In the event where the two or more observations are tied, SPSS assigns the average rank to each tied observation to calculate the mean rank for each sample. After ranking all values, the Kruskal-Wallis H statistic is calculated via:

$$H = \left[\frac{12}{n(n+1)}\sum_{j=1}^{c}\frac{T_j^2}{n_j}\right] - 3(n+1) \quad (1)$$

where n = sum of sample sizes for all samples, c = number of samples, T_j = sum of ranks in the jth sample, n_j = size of the jth sample.

The results for each test are shown in Table 1, with how each question performed, their mean rank which has a maximum value of 60 given our number of participants (the mean of the ranks assigned to the data since it is more appropriate to use ranks over values to prevent testing being affected by the presence of outliers), median for each group, standard deviation, Kruskal-Wallis H statistic, and the number of participants that used the method is shown in the final column. If the question has an assumed significance figure (p-value) that is less than the alpha value (significance level of) 0.05, it can be considered statistically significant, as indicated by an asterisk in the p-value column, else it is not statistically significant.

After conducting the experiment, the Kruskal-Wallis test tells us that 15 of the 27 questions were found to have a statistical significance between them. The remaining questions were not considered statistically significant; however, some conclusions may apprehensively be drawn from them. We then ran the post-Hoc test using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons on the statistically significant questions to determine which authentication methods specifically were significant to one another seeing if there were any pairwise comparisons between the methods. The reason we chose this test was because we have a small subset of all possible pairs.

I believe the authentication method is not easily hacked?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF (authentication factor) scores between the voice (3.00) and finger (4.00) (p = 0.001) but not with PIN (3.00), token (4.00) or any other group combination (Table 2).

In regard to mean ranking fingerprint ranked the highest, followed by token, PIN and finally voice (Figure 2). Users perhaps considered fingerprint the hardest to hack due to it relying on inheritance-based authentication that only the user possess. Despite this however, voice another inheritance-based authentication ranked last, hence users felt as though voice was easily hacked, this might be because users believe the sensors can be easily spoofed due to issues with background noise or voice changing. Meanwhile both token and PIN ranked in between. This is likely because users are used to both these methods. There was a statistical significance found between the methods voice and fingerprint, hence it can be concluded that users consider fingerprint harder to hack than voice.

Table 1. Kruskal Wallis test results

Question		Sample Size (n)	Method	Mean Rank	Median	Std. Dev.	H	p-Value
Availability	Method is available when needed?	15	Voice	28.60	5	0.497	2.248	0.523
		15	PIN	33.10	5			
		15	Fingerprint	27.20	5			
		15	Token	33.10	5			
	The authentication method is widespread?	15	Voice	29.67	4	0.841	3.060	0.383
		15	PIN	36.50	5			
		15	Fingerprint	26.77	4			
		15	Token	29.07	4			
	A common example of authentication techniques?	15	Voice	25.40	4	0.848	3.704	0.295
		15	PIN	36.07	5			
		15	Fingerprint	29.40	5			
		15	Token	31.13	5			
Security	I believe the authentication method is not easily hacked?	15	Voice	18.07	3	1.136	15.199	0.002*
		15	PIN	30.10	3			
		15	Fingerprint	41.63	4			
		15	Token	32.20	4			
	Method is not easily tampered with?	15	Voice	21.63	3	1.031	6.547	0.088
		15	PIN	30.33	4			
		15	Fingerprint	35.87	4			
		15	Token	34.17	4			
	Method is able to differentiate me?	15	Voice	28.33	4	1.331	19.586	$p < 0.001$*
		15	PIN	23.97	4			
		15	Fingerprint	46.67	5			
		15	Token	23.03	3			
Usability	Learning to use the authentication method easy?	15	Voice	27.63	5	0.431	4.336	0.227
		15	PIN	33.53	5			
		15	Fingerprint	27.30	5			
		15	Token	33.53	5			
	I found the authentication method easy to use?	15	Voice	27.90	5	0.628	6.553	0.088
		15	PIN	36.00	5			
		15	Fingerprint	25.97	5			
		15	Token	32.13	5			
	Accessible to different needs?	15	Voice	33.40	4	1.017	1.392	0.707
		15	PIN	30.63	4			
		15	Fingerprint	31.47	4			
		15	Token	26.50	4			

continued on following page

Table 1. Continued

Question		Sample Size (n)	Method	Mean Rank	Median	Std. Dev.	H	p-Value
Privacy	Protects my privacy from others?	15	Voice	15.50	3	1.030	22.332	$p < 0.001$*
		15	PIN	37.13	5			
		15	Fingerprint	42.07	5			
		15	Token	27.30	4			
	Prevents others from seeing my data?	15	Voice	19.50	3	1.039	9.503	0.023*
		15	PIN	35.77	4			
		15	Fingerprint	35.60	4			
		15	Token	31.13	4			
	Method allows me to remain anonymous?	15	Voice	22.80	3	1.307	8.198	0.042*
		15	PIN	38.23	4			
		15	Fingerprint	34.87	4			
		15	Token	26.10	3			
Reliability	Authentication functioned as I expected it to?	15	Voice	29.10	5	0.504	3.469	0.325
		15	PIN	35.00	5			
		15	Fingerprint	29.10	5			
		15	Token	29.10	5			
	Method performed the same each time?	15	Voice	24.83	4	0.813	11.205	0.011*
		15	PIN	40.50	5			
		15	Fingerprint	25.43	5			
		15	Token	31.23	5			
	Will continue to perform as expected in further uses?	15	Voice	25.60	4	0.701	16.310	0.001*
		15	PIN	41.50	5			
		15	Fingerprint	21.20	4			
		15	Token	33.70	5			
Experience	I used the authentication method many times before?	15	Voice	15.27	2	1.570	20.773	$p < 0.001$*
		15	PIN	39.53	5			
		15	Fingerprint	32.57	5			
		15	Token	34.63	5			
	I use similar authentication methods often?	15	Voice	19.43	2	1.448	12.370	0.006*
		15	PIN	39.53	5			
		15	Fingerprint	19.43	4			
		15	Token	33.57	5			
	I use different authentication methods often?	15	Voice	34.80	5	1.247	5.056	0.168
		15	PIN	35.27	5			
		15	Fingerprint	28.10	4			
		15	Token	23.83	4			

continued on following page

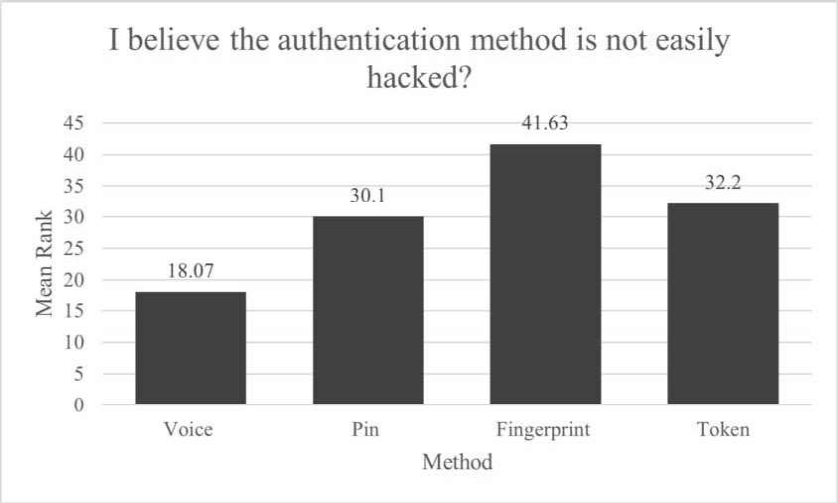
Table 1. Continued

Question		Sample Size (n)	Method	Mean Rank	Median	Std. Dev.	H	p-Value
Verification	Feedback that authentication has processed correctly?	15	Voice	21.17	4	0.616	8.376	0.039*
		15	PIN	37.17	5			
		15	Fingerprint	32.37	5			
		15	Token	31.30	5			
	Feedback when a type of error has occurred?	15	Voice	20.70	3	1.145	18.475	$p < 0.001$*
		15	PIN	40.40	4			
		15	Fingerprint	39.27	4			
		15	Token	21.63	3			
	Good feedback that it has been set up correctly?	15	Voice	33.30	5	0.930	8.064	0.045*
		15	PIN	33.07	5			
		15	Fingerprint	35.03	5			
		15	Token	20.60	4			
Knowledge	I understand how the method works?	15	Voice	26.50	4	0.780	3.982	0.263
		15	PIN	36.80	5			
		15	Fingerprint	27.40	4			
		15	Token	31.30	5			
	I have a good understanding of the authentication process works?	15	Voice	23.27	4	0.914	9.184	0.027*
		15	PIN	37.87	5			
		15	Fingerprint	25.43	4			
		15	Token	35.43	5			
	Understanding of why authentication method is used?	15	Voice	23.30	4	0.623	6.022	0.111
		15	PIN	36.10	5			
		15	Fingerprint	29.43	5			
		15	Token	33.17	5			
Recommendation	I have heard others have good experiences with the authentication	15	Voice	19.07	3	1.152	12.951	0.005*
		15	PIN	37.77	5			
		15	Fingerprint	37.53	5			
		15	Token	27.63	4			
	I have heard others have bad experiences with the authentication	15	Voice	30.37	2	1.239	1.572	0.666
		15	PIN	29.77	3			
		15	Fingerprint	34.77	3			
		15	Token	27.10	2			
	The authentication method has a good reputation?	15	Voice	20.10	3	0.933	9.053	0.029*
		15	PIN	34.40	4			
		15	Fingerprint	37.00	4			
		15	Token	30.50	4			

Table 2. Post hoc 'is not easily hacked'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-PIN	12.033	6.087	1.977	.048	.288
Voice-Token	-14.133	6.087	-2.322	.020	.121
Voice-Finger	23.567	6.087	3.871	.000	.001*
PIN-Token	-2.100	6.087	-.345	.730	1.000
PIN-Finger	11.533	6.087	1.895	.058	.349
Token-Finger	9.433	6.087	1.550	.121	.727

Figure 2. Mean rank of method not being easily hacked



I believe the authentication method is able to differentiate me from others?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the finger (5) and voice (3) ($p = 0.016$), finger and PIN (4) ($p = 0.001$) and finger and token (4) ($p = 0.001$) but not with any other group combination (Table 3).

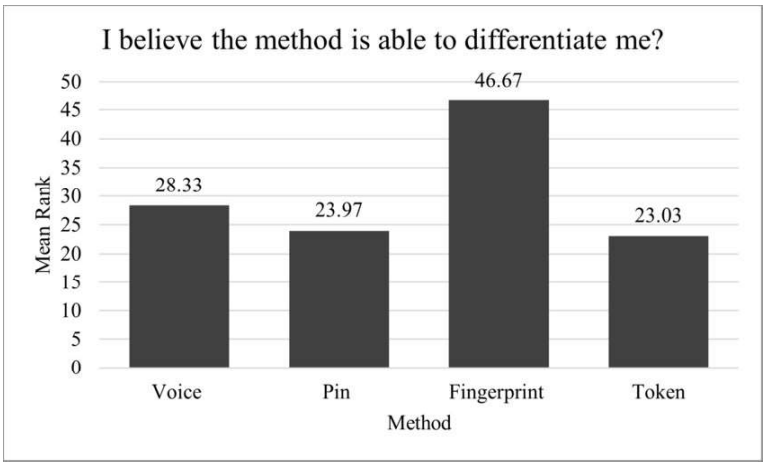
In regard to mean ranking fingerprint again ranked the highest, followed by voice, PIN and finally token (Figure 3). Both the inheritance-based authentication methods ranked the highest, likely because users consider biometrics exclusive to just the user, whereas knowledge/owner-based methods ranked lower, likely because if another user used those methods the system would not be able to tell the difference. There was a statistical significance between fingerprint and voice, fingerprint and PIN and fingerprint and token, meaning that it can be concluded that fingerprint was considered the best at being able to differentiate users from others.

I believe the authentication method protects my privacy from others?

Table 3. Post hoc 'differentiate me from others'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p value	Adj. p value
Voice-PIN	-4.367	6.101	-0.716	.474	1.000
Voice-Token	5.300	6.101	0.869	.385	1.000
Voice-Finger	18.333	6.101	3.005	.003	0.016*
PIN-Token	0.933	6.101	0.153	.878	1.000
PIN-Finger	22.700	6.101	3.721	.000	0.001*
Token-Finger	23.633	6.101	3.874	.000	0.001*

Figure 3. Mean rank of method differentiating



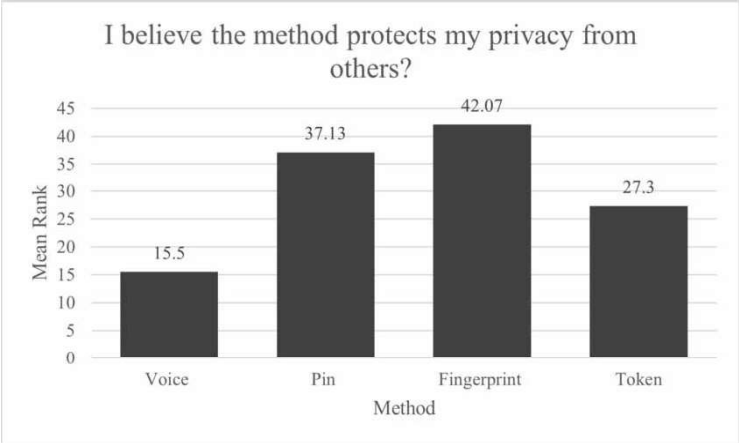
Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (3) and PIN (5) ($p = 0.002$) and voice and finger (5) ($p < 0.001$) but not with token (4) or any other group combination (Table 4).

For mean ranking fingerprint ranked the highest, followed by PIN, then token and finally voice (Figure 4). Users found fingerprint to be the most likely to protect their privacy from others perhaps

Table 4. Post hoc 'protects my privacy from others'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Token	-11.800	6.082	-1.940	.052	.314
Voice-PIN	21.633	6.082	3.557	.000	.002*
Voice-Finger	26.567	6.082	4.368	.000	.000*
PIN-Token	9.833	6.082	1.617	.106	.636
Token-Finger	14.767	6.082	2.428	.015	.091
PIN-Finger	4.933	6.082	.811	.417	1.000

Figure 4. Mean rank of protecting privacy from others



because people believe biometrics are extremely hard to spoof. However, voice ranked the lowest, meaning users do not believe that voice will protect their privacy because they presumably believe that voice could be more easily spoofed compared to fingerprints. Both PIN and tokens ranked in the middle, less than fingerprint but more than voice, likely because they are used to those authentication methods. There was a statistical significance between voice with PIN and voice with fingerprint. Hence, it can be assumed that users believe PIN and fingerprint to be more likely to protect their user's privacy compared to voice.

I believe the authentication method prevents others from seeing the contents of my data?

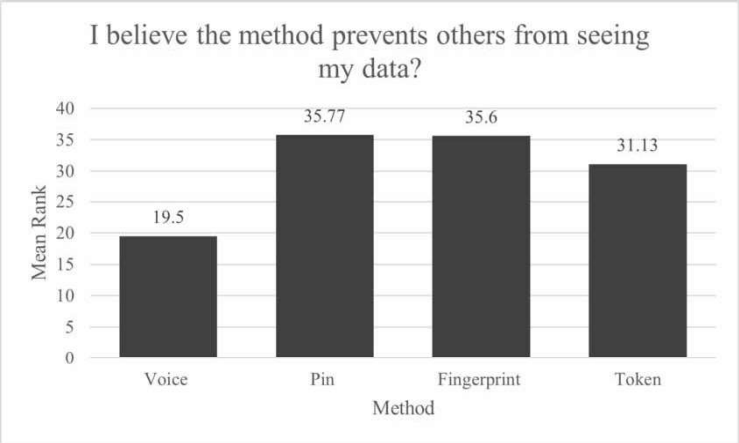
Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (3) and PIN (4) ($p = 0.044$) and voice and finger (4) ($p = 0.048$) but not with token (4) or any other group combination (Table 5).

For mean ranking PIN ranked the highest, followed by fingerprint, token and finally voice (Figure 5). Voice ranked the lowest again likely because users believe that it can be easily spoofed and hence does not protect the contents of their data. However, PIN ranked slightly higher than fingerprint. This is likely because users are most used to PIN and hence had more faith it would prevent others from seeing their data. There was a statistical significance between voice with PIN and voice with

Table 5. Post hoc 'prevents others from seeing my data'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Token	-11.633	6.071	-1.916	.055	.332
Voice-Finger	16.100	6.071	2.652	.008	.048*
Voice-PIN	16.267	6.071	2.679	.007	.044*
PIN-Token	4.633	6.071	.763	.445	1.000
Token-Finger	4.467	6.071	.736	.462	1.000
PIN-Finger	-.167	6.071	-.027	.978	1.000

Figure 5. Mean rank preventing others from seeing data



fingerprint. Hence, it can be assumed that users believe PIN and fingerprint to be more likely to prevent others from seeing their data compared to voice.

I believe the authentication method allows me to remain anonymous?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed no statistically significant differences in median AF scores between any pairwise comparisons (Table 6). The median scores were voice (3), PIN (4), finger (4) and token (3).

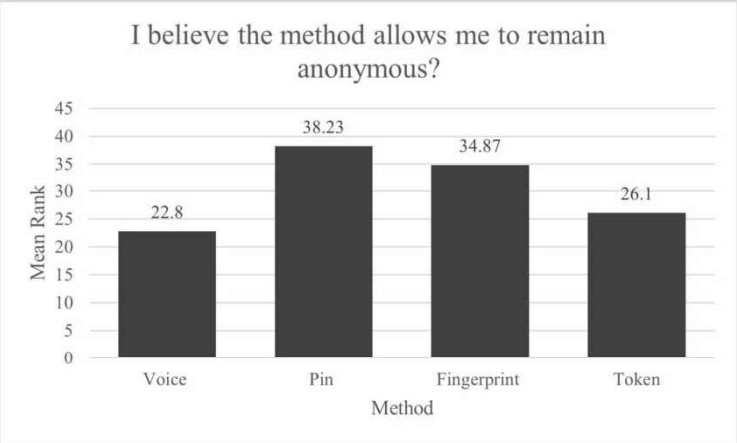
When concerned with mean ranking, PIN ranked the highest, followed by fingerprint, token and finally voice (Figure 6). Users ranked PIN the highest likely because users do not have to give any personal info or other accounts to utilise a PIN, whereas other methods like token usually require you to link some other device or account, hence users felt they remained less anonymous. Both biometrics also ranked lower than PIN likely because users must use their personal inheritance features. The results were found to be statistically significant however there were no specific groups that were statistically significant to one another.

The authentication method performed the same each time?

Table 6. Post hoc 'allows me to remain anonymous'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Token	-3.300	6.199	-.532	.594	1.000
Voice-Finger	12.067	6.199	1.947	.052	.310
Voice-PIN	15.433	6.199	2.490	.013	.077
Token-Finger	8.767	6.199	1.414	.157	.944
PIN-Token	12.133	6.199	1.957	.050	.302
PIN-Finger	-3.367	6.199	-.543	.587	1.000

Figure 6. Mean rank of allowing to remain anonymous



Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (4) and PIN (5) ($p = 0.019$) and PIN and finger (5) ($p = 0.028$) but not with token (5) or any other group combination (Table 7).

PIN ranked the highest followed by token, fingerprint and finally voice (Figure 7). Both authentication methods that use a keyboard to input ranked higher than the inheritance methods. This is likely because users have the most control over the input by inputting it themselves, whereas with inheritance-based methods it relies entirely on the sensor being able to recognise the users' input. There was a statistical significance between voice and PIN as well as between PIN and fingerprint. Hence, we can conclude that users found PIN to perform more consistently than voice and fingerprint authentication methods.

The authentication method will continue to perform as expected in further uses?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (4) and PIN (5) ($p = 0.021$) and PIN and finger (4) ($p = 0.001$) but not with token (5) or any other group combination (Table 8).

Table 7. Post hoc 'performed the same each time'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Finger	0.600	5.316	.113	.910	1.000
Voice-Token	-6.400	5.316	-1.204	.229	1.000
Voice-PIN	15.667	5.316	2.947	.003	.019*
PIN-Finger	-15.067	5.316	-2.834	.005	.028*
Token-Finger	-5.800	5.316	-1.091	.275	1.000
PIN-Token	9.367	5.316	1.743	.081	.488

Figure 7. Mean rank of method performing same each time

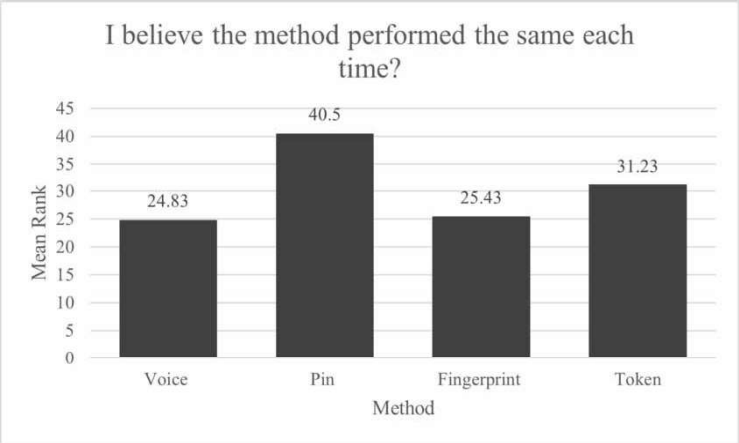


Table 8. Post hoc 'perform as expected in further uses'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Finger	-4.400	5.445	-.808	.419	1.000
Token-Finger	-12.500	5.445	-2.296	.022	.130
PIN-Finger	-20.300	5.445	-3.728	.000	.001*
Voice-Token	-8.100	5.445	-1.488	.137	.821
Voice-PIN	15.900	5.445	2.920	.003	.021*
PIN-Token	7.800	5.445	1.433	.152	.912

For mean ranking, PIN ranked the highest, followed by token, fingerprint and finally voice (Figure 8). Likewise, much like the above concerns about the methods performing the same each time, the two methods that rely on the users to input the authentication themselves, users considered to be more reliable for further uses. Whereas the inheritance-based methods ranked lower, perhaps due to them relying on the sensor having to recognising the user. There was found to be a statistical significance again between PIN and voice as well as PIN and fingerprint. Therefore, it can be concluded that users expect PIN to perform more consistently than both fingerprint and voice.

I have used the authentication method many times before?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (2) and finger (5) ($p = 0.014$), voice and token (5) ($p = 0.004$) and voice and PIN (5) ($p < 0.001$) but not with any other group combination (Table 9).

PIN ranked the highest for mean ranking, followed by token, fingerprint and finally voice (Figure 9). Unsurprisingly traditional means of authentication such as PIN and token ranked the highest since they are the most common means of authentication. Voice ranked the last considering the authentication method is reasonably new. There was found to be a statistical difference between

Figure 8 Mean rank of method perform as expected with further use

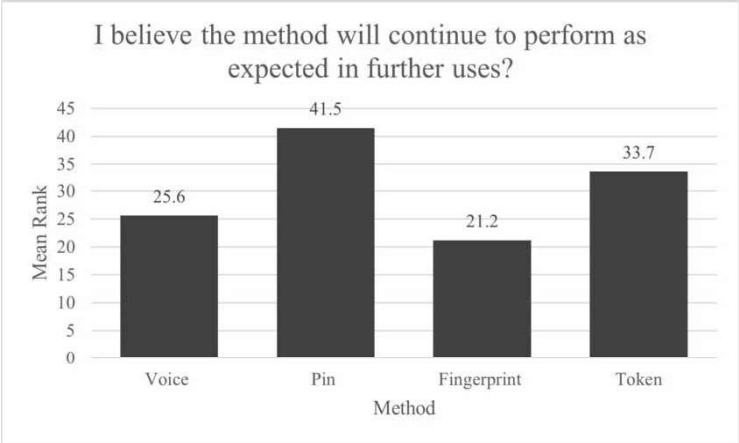
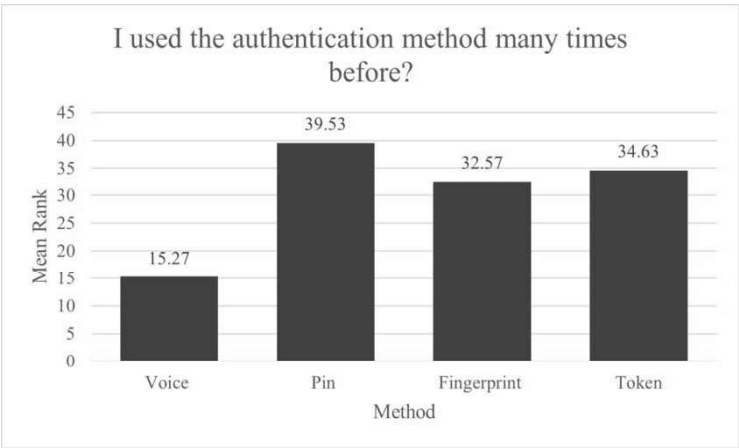


Table 9. Post hoc 'used the method many times before'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Finger	17.300	5.679	3.046	.002	.014*
Voice-Token	-19.367	5.679	-3.410	.001	.004*
Voice-PIN	24.267	5.679	4.273	.000	.000*
Token-Finger	-2.067	5.679	-.364	.716	1.000
PIN-Finger	-6.967	5.679	-1.227	.220	1.000
PIN-Token	4.900	5.679	.863	.388	1.000

Figure 9. Mean rank of used method many times before



voice and PIN, voice and fingerprint and voice and token. Therefore, it can be concluded that users have all used PINs, fingerprints, and tokens more than voice authentication.

I use similar authentication methods often?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (2) and PIN (5) ($p = 0.004$) but not with finger (4), token (5) or any other group combination (Table 10).

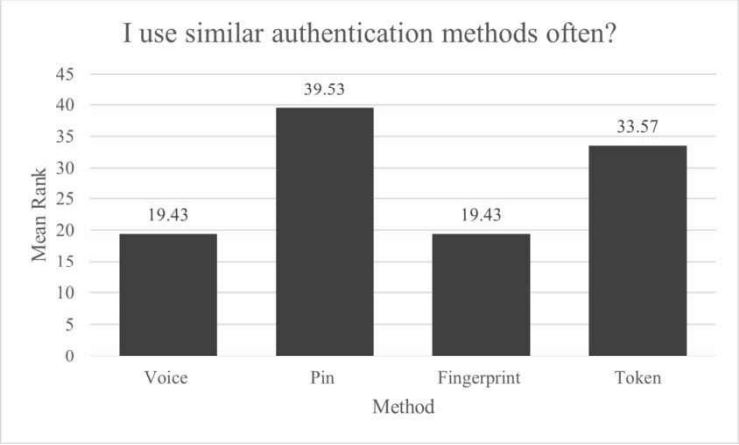
Once again PIN ranked the highest for mean ranking, followed by token and fingerprint and voice ranked last (Table 10). PIN and token ranking high is likely because knowledge-based authentication and ownership-based authentication are more common whereas the biometric methods ranked much lower, likely because they are less common. There was found to be a statistical difference between voice and PIN. Therefore, it can be concluded that users use methods similar to PIN's far more often than they use methods similar to voice authentication.

I believe the authentication method offers good feedback that my authentication has processed correctly?

Table 10. Post hoc 'used similar methods often'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Finger	10.033	5.890	1.704	.088	.531
Voice-Token	-14.133	5.890	-2.400	.016	.098
Voice-PIN	20.100	5.890	3.413	.001	.004*
Token-Finger	-4.100	5.890	-.696	.486	1.000
PIN-Finger	-10.067	5.890	-1.709	.087	.525
PIN-Token	5.967	5.890	1.013	.311	1.000

Figure 10. Mean rank of use similar methods often



Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (4) and PIN (5) ($p = 0.030$) but not with finger (5), token (5) or any other group combination (Table 11).

For mean ranking, PIN ranked the highest, followed by fingerprint, token and finally voice (Figure 11). For mean ranking PIN ranked the highest, given users input the numbers themselves and are immediately singed-in providing they gave the correct PIN. Meanwhile, an authentication method such as fingerprint or voice can be inputted, yet it does not always sign the user in as what the sensor saw/heard did not exactly match. There was found to be a statistical difference between voice and PIN. Therefore, it can be concluded that users consider PIN to offer better feedback than voice that their authentication has processed correctly.

I believe the authentication method offers good feedback when some type of error has occurred?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (3) and finger (4) ($p = 0.015$), voice and PIN (4) ($p = 0.008$), PIN and token (3) ($p = 0.014$) and token and finger ($p = 0.025$) but not with any other group combination (Table 12).

Table 11. Post hoc 'authentication processed correctly'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Token	-10.133	5.692	-1.780	.075	.450
Voice-Finger	11.200	5.692	1.968	.049	.295
Voice-PIN	16.000	5.692	2.811	.005	.030*
Token-Finger	1.067	5.692	.187	.851	1.000
PIN-Token	5.867	5.692	1.031	.303	1.000
PIN-Finger	-4.800	5.692	-.843	.399	1.000

Figure 11. Mean rank of method offering feedback authentication processed correctly

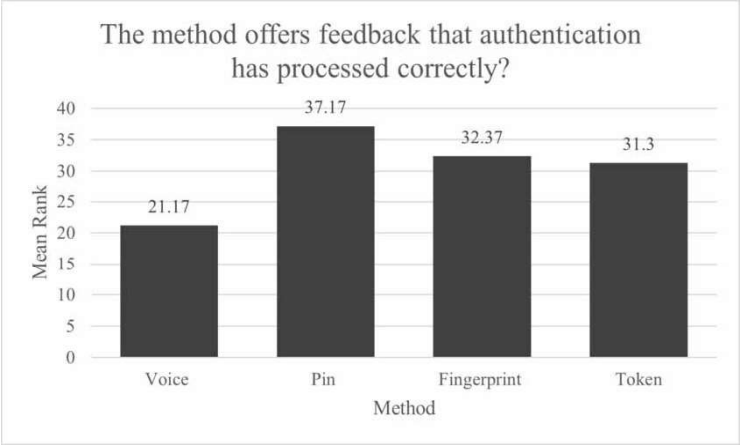


Table 12. Post hoc 'feedback when error has occurred'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Token	-.933	6.151	-.152	.879	1.000
Voice-Finger	18.567	6.151	3.018	.003	.015*
Voice-PIN	19.700	6.151	3.203	.001	.008*
Token-Finger	17.633	6.151	2.867	.004	.025*
PIN-Token	18.767	6.151	3.051	.002	.014*
PIN-Finger	-1.133	6.151	-.184	.854	1.000

In regard to mean ranking PIN ranked the highest, followed by fingerprint, token and finally voice (Figure 12). PIN again ranked the highest likely because when a user inputs the PIN, they are immediately either signed in or told the pin code was incorrect and a user knows a digit was wrong. Whereas a method such as voice will not sign in but there are a lot more factors as to why the voice print did not match, such as the voice itself or background noise. Token likewise ranked low as again there can be multiple reasons why the token was wrong i.e., had it been inputted wrong or had it expired. There was found to be a statistical difference between voice and PIN, voice and fingerprint, PIN and token and fingerprint and token. Therefore, it can be concluded that users consider PIN and fingerprint to offer much better feedback when a type of error has occurred compared to voice and tokens.

I believe the authentication method offers good feedback that it has been set up correctly?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed no statistically significant differences in median AF scores between any pairwise comparisons (Table 13). The median scores were voice (5), PIN (5), finger (5) and token (4).

Fingerprint ranked the highest for mean ranking, followed by voice, PIN and finally token (Figure 13). Fingerprint ranked the highest as in the setup of the method, it usually guides the user through

Figure 12. Mean rank of method offering feedback when error occurs

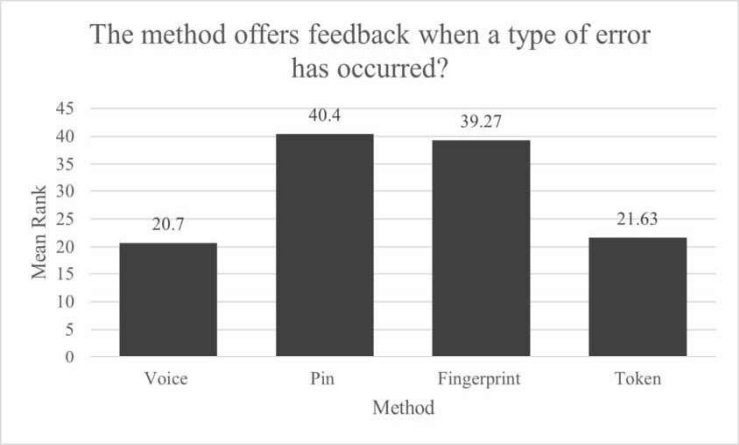
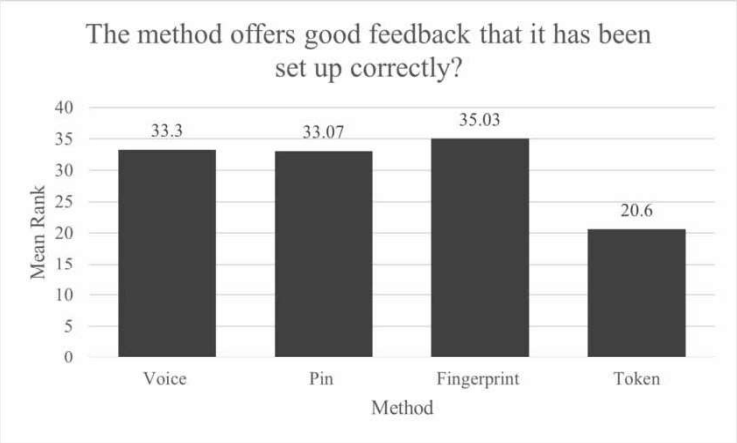


Table 13. Post hoc 'feedback when set up correctly'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
PIN-Token	12.467	5.743	2.171	.030	.180
Voice-Token	12.700	5.743	2.211	.027	.162
Token-Finger	14.433	5.743	2.513	.012	.072
Voice-PIN	-.233	5.743	-.041	.968	1.000
PIN-Finger	1.967	5.743	.342	.732	1.000
Voice-Finger	1.733	5.743	.302	.763	1.000

Figure 13. Mean rank of feedback when set up correctly



building up their print slowly, likewise voice similarly builds up the print over a few recordings. Token meanwhile offers less feedback during setup and usually requires a test to see that it has been set up correctly. The results were found to be statistically significant however there were no specific groups that were statistically significant to one another.

I have a good understanding of how the authentication process works?

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed no statistically significant differences in median AF scores between any pairwise comparisons (Table 14). The median scores were voice (4), PIN (5), finger (5) and token (5).

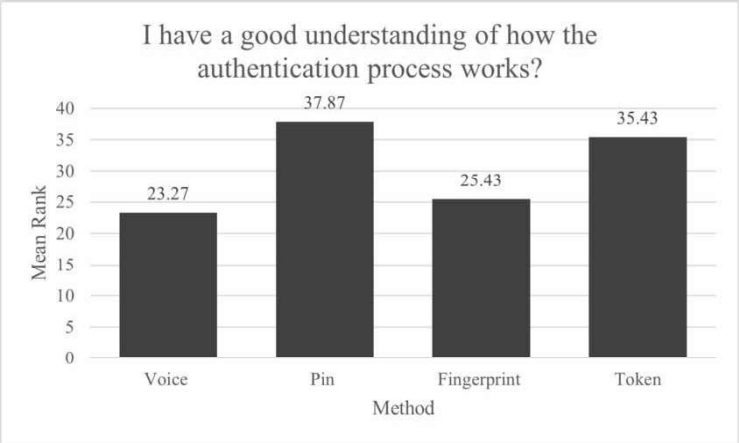
For meaning ranking, PIN ranked the highest, followed by token, then fingerprint and finally voice (Figure 14). PIN and token ranked the highest likely due to them being more commonly used, hence users are much more likely to have an understanding of how the authentication process works. Whereas the less commonly used methods such as voice and fingerprint ranked lower, likely because users were less used to those methods. The results were found to be statistically significant however there were no specific groups that were statistically significant to one another.

I have heard others have good experiences with the authentication method?

Table 14. Post hoc 'understanding of how process works'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Finger	2.167	5.840	.371	.711	1.000
Voice-Token	-12.167	5.840	-2.083	.037	.223
Voice-PIN	14.600	5.840	2.500	.012	.074
Token-Finger	-10.000	5.840	-1.712	.087	.521
PIN-Finger	-12.433	5.840	-2.129	.033	.199
PIN-Token	2.433	5.840	.417	.677	1.000

Figure 14. Mean rank of user understanding authentication process



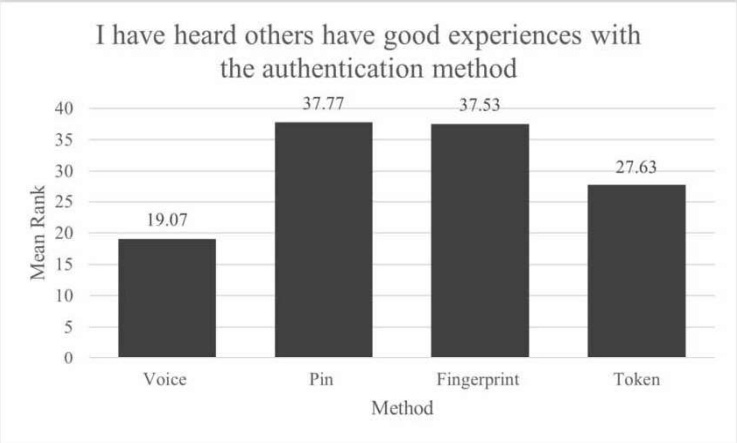
Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (3) and finger (5) ($p = 0.015$) and voice and PIN (5) ($p = 0.013$) but not with token (4) or any other group combination (Table 15).

For mean ranking PIN ranked the highest, followed by fingerprint, token and finally voice (Figure 15). Notably voice ranked the lowest here, having also been the method with the least prior

Table 15. Post hoc 'heard good experience'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Token	-8.567	6.103	-1.404	.160	.963
Voice-Finger	18.467	6.103	3.026	.002	.015*
Voice-PIN	18.700	6.103	3.064	.002	.013*
Token-Finger	9.900	6.103	1.622	.105	.629
PIN-Token	10.133	6.103	1.660	.097	.581
PIN-Finger	-.233	6.103	-.038	.970	1.000

Figure 15. Mean rank of good experience heard from others



usage by users. Users not hearing others have good experiences with the method could be why many users have not used voice. Meanwhile PIN and fingerprint both ranked much higher, surprisingly fingerprint ranked high despite only ranking third for users having used the method before. There was found to be a statistical difference between voice and PIN, voice, and fingerprint. Therefore, it can be concluded that users find they have heard others have a better experience with PIN and fingerprint compared to voice.

The authentication method has a good reputation?

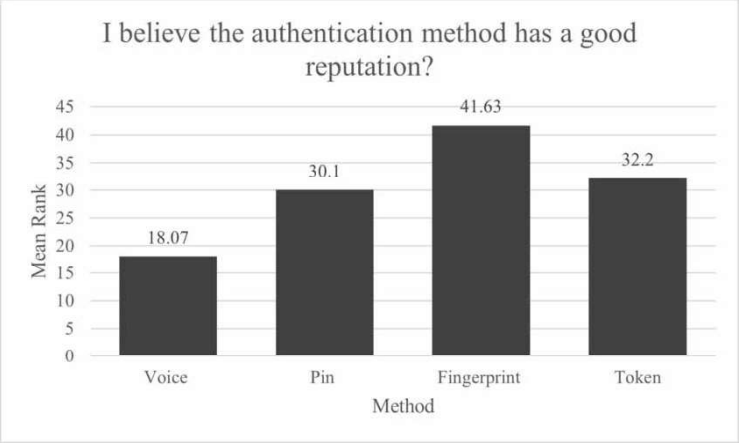
Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. This post hoc analysis revealed statistically significant differences in median AF scores between the voice (3) and finger (4) ($p = 0.031$) but not with finger (4), token (4) or any other group combination (Table 16).

Fingerprint ranked the highest for mean ranking, followed by PIN, token and finally voice (Figure 16). Voice again ranked lower than other methods, indicating that many users would not want to use the authentication method as they feel as though they do not have a good reputation. Comparably the other biometric method fingerprint, ranked the highest in terms of reputation indicating many users believe fingerprint to have an excellent reputation. There was found to be a statistical difference between

Table 16. Post hoc 'method has good reputation'

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	p Value	Adj. p Value
Voice-Token	-10.400	6.049	-1.719	.086	.513
Voice-PIN	14.300	6.049	2.364	.018	.108
Voice-Finger	16.900	6.049	2.794	.005	.031*
PIN-Token	3.900	6.049	.645	.519	1.000
Token-Finger	6.500	6.049	1.075	.283	1.000
PIN-Finger	2.600	6.049	.430	.667	1.000

Figure 16. Mean rank of authentication method having good reputation



voice and fingerprint. Therefore, it can be concluded that users consider fingerprint authentication to have a much better reputation compared to voice.

Of these, the main trends suggested that users were more likely to trust PIN out of the four methods, given it ranked the highest across categories such as privacy, reliability, experience, verification, knowledge, and recommendation. Voice, meanwhile, was the method that ranked the lowest most often, indicating that users would be unlikely to trust and therefore utilise voice biometric authentication over methods they had more trust with.

5. CONCLUSION

In this study, the expanded trust model (Hoffman et al., 2006) is used to measure how users trust authentication methods when accessing IoT technologies between four different authentication methods: PIN, Tokens, Fingerprints, and Voice. The purpose was to discern if users would be willing to utilise voice biometric authentication by comparing the levels of trust, they have with the method compared to traditional authentication means. To achieve this, we derived a realistic trust evaluation model that incorporates privacy, reliability, security, usability, safety, and availability factors into a trust vector for a flexible measurement of trust in the user accessing the technology. Using a Kruskal-Wallis H test and the post-hoc test, the collected data were examined to find if there is any statistical significance. Based on the study results, around half of the results were found to be statistically significant compared to one another, with the main trends suggesting that users were more likely to trust passwords out of the four authentication methods. This was the case since it ranked the highest across categories such as privacy, reliability, experience, verification, knowledge, and recommendation. Meanwhile, the voice biometric method was most often ranked the lowest, indicating that users would be unlikely to trust it compared to the other three authentication methods.

Since users' perceptions change with time through continued use of technology, allowing perceptions, opinions, and trust to change. There are several practical applications to this study. From the trust model itself, the results highlight the particular weaknesses users perceive voice biometrics to possess in comparison to other methods of authentication. As such, the future development of voice biometric systems should consider implementing systems to address each of the aspects of trust with which the users expressed concerns. Another practical application of this study is that measuring a user's trust in technology can potentially reduce risk. For example, if consumers lack trust in a technology, they are more likely to use it with caution or abandon it altogether. This can

expose developers to significant financial hazards. Technology developers can reduce the risk of creating technology that fails to gain traction or is abandoned by users by gaining a comprehension of user opinions based on the proposed trust metrics in this study. Subsequently, developers can identify areas in which they can differentiate their technology and earn the users' trust by gaining an understanding of the users' perceptions of trust.

REFERENCES

Arora, S., & Bhatia, M. (2021) Challenges and opportunities in biometric security: A survey. Information Security Journal: A Global Perspective, 1-21.

Bala, N., Gupta, R., & Kumar, A. (2021) Multimodal biometric system based on fusion techniques: A review. Information Security Journal: A Global Perspective, 1-49.

Chen, W., Hancke, G. P., Mayes, K. E., Lien, Y., & Chiu, J. H. (2010) Using 3G network components to enable NFC mobile transactions and authentication. *2010 IEEE International Conference on Progress in Informatics and Computing*. doi:10.1109/PIC.2010.5687587

Chokhani, S. (2004) Knowledge based authentication (KBA) metrics. *KBA Symposium-Knowledge Based Authentication: Is It Quantifiable.*

Choong, Y., Theofanos, M. F., Renaud, K., & Prior, S. (2019). "Passwords protect my stuff"—A study of children's password practices. *Journal of Cybersecurity*, *5*(1), tyz015. doi:10.1093/cybsec/tyz015 PMID:33042580

Chrobok Mateusz. (2020). *Physical biometrics vs behavioral biometrics*. Available from https://www.buguroo.com/en/blog/physical-biometrics-vs-behavioral-biometrics

Cruz, J., Mishra, P., & Bhunia, S. (2019) The metric matters: The art of measuring trust in EIectronics. *2019 56th ACM/IEEE Design Automation Conference (DAC).*

Dietz, G. (2011). Going back to the source: Why do people trust each other? *Journal of Trust Research*, *1*(2), 215–222. doi:10.1080/21515581.2011.603514

Fu, K. (2015) *Knowledge-based authentication (kba)*. Available from https://www.hsgac.senate.gov/imo/media/doc/Testimony-Fu-2015-06-02-REVISED%2021.pdf

Hoffman, L. J., Lawson-Jenkins, K., & Blum, J. (2006). Trust beyond security: An expanded trust model. *Communications of the ACM*, *49*(7), 95–101. doi:10.1145/1139922.1139924

ISO/IEC. (2007). *Identification cards — integrated circuit cards — part 2: Cards with contacts — dimensions and location of the contacts*. ISO.

Jones, R. (2018) *Voice recognition: Is it really as secure as it sounds?* Available from https://www.theguardian.com/money/2018/sep/22/voice-recognition-is-it-really-as-secure-as-it-sounds

Korshunov, P., & Marcel, S. (2017). Impact of score fusion on voice biometrics and presentation attack detection in cross-database evaluations. *IEEE Journal of Selected Topics in Signal Processing*, *11*(4), 695–705. doi:10.1109/JSTSP.2017.2692389

Krawczyk, S., & Jain, A. K. (2005). Securing electronic medical records using biometric authentication. *International Conference on Audio-and Video-Based Biometric Person Authentication.* doi:10.1007/11527923_115

Krom, G. (1994). Consistency and reliability of voice quality ratings for different types of speech fragments. *Journal of Speech, Language, and Hearing Research: JSLHR*, *37*(5), 985–1000. doi:10.1044/jshr.3705.985 PMID:7823566

Kunyu, P., Jiande, Z., & Jing, Y. (2009). An identity authentication system based on mobile phone token. *2009 IEEE International Conference on Network Infrastructure and Digital Content.* doi:10.1109/ICNIDC.2009.5360974

Kuo, C., & Lo, M. (1999). Secure Open Smart Card Architecture. Academic Press.

Lavrentyeva, G., Novoselov, S., Malykh, E., Kozlov, A., Kudashev, O., & Shchemelinin, V. (2017). *Audio replay attack detection with deep learning frameworks*. Interspeech. doi:10.21437/Interspeech.2017-360

Lim, C. H., & Kwon, T. (2006) Strong and robust RFID authentication enabling perfect ownership transfer. *International Conference on Information and Communications Security.* doi:10.1007/11935308_1

Marcel, S., Nixon, M. S., Fierrez, J., & Evans, N. (2019). *Handbook of biometric anti-spoofing: Presentation attack detection*. Springer. doi:10.1007/978-3-319-92627-8

McKight, P. E., & Najab, J. (2010). *Kruskal-wallis test*. The Corsini Encyclopedia of Psychology.

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

Nuance. (2018). *Nuance unveils AI-powered virtual assistant solution designed for healthcare providers.* Available from https://www.nuance.com/en-gb/about-us/newsroom/press-releases/nuance-unveils-AI-Powered-solution-for-healthcare.html

Nuance. (n.d.). *Barclays improves their customer experience.* Available from https://www.nuance.com/omni-channel-customer-engagement/case-studies/barclays.html

Ortega-Garcia, J., Bigun, J., Reynolds, D., & Gonzalez-Rodriguez, J. (2004). Authentication gets personal with biometrics. *IEEE Signal Processing Magazine*, *21*(2), 50–62. doi:10.1109/MSP.2004.1276113

Patent, V., & Searle, R. H. (2019). Qualitative meta-analysis of propensity to trust measurement. *Journal of Trust Research*, *9*(2), 136–163. doi:10.1080/21515581.2019.1675074

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, *49*(1), 95–112. doi:10.1037/0022-3514.49.1.95 PMID:11474726

Rigas, I., Abdulin, E., & Komogortsev, O. (2016). Towards a multi-source fusion approach for eye movement-driven recognition. *Information Fusion*, *32*, 13–25. doi:10.1016/j.inffus.2015.08.003

Shablygin, E., Zakharov, V., Bolotov, O., & Scace, E. (2013). *Token management.* United States *Patent* no. US8555079, B2.

Tupman, A. (2018). *5 reasons why voice biometrics is A game changer.* Available from https://www.conn3ct.com/blog/five-reasons-why-voice-biometrics-is-a-game-changer

Usman, A. B., & Gutierrez, J. (2018). Toward trust based protocols in a pervasive and mobile computing environment: A survey. *Ad Hoc Networks*, *81*, 143–159. doi:10.1016/j.adhoc.2018.07.009

Usman, A. B., & Gutierrez, J. (2019). DATM: A dynamic attribute trust model for efficient collaborative routing. *Annals of Operations Research*, *277*(2), 293–310. doi:10.1007/s10479-018-2864-5

Usman, A. B., Gutierrez, J. A., & Bichi, A. B. (2019). Secure routing protocols using trust-based mechanisms in the internet of things for smart city environment challenges and future trends. In Anonymous secure cyber-physical systems for smart cities. IGI Global. doi:10.4018/978-1-5225-7189-6.ch005

Yang, W., Wang, S., Sahri, N. M., Karie, N. M., Ahmed, M., & Valli, C. (2021). Biometrics for internet-of-things security: A review. *Sensors (Basel)*, *21*(18), 6163. doi:10.3390/s21186163 PMID:34577370

Alec Wells is a postgraduate PhD student at York St John University in the UK. He received a Master's by Research degree also at York St John University in the UK. His research interests include cyber security, authentication, biometrics and IoT.

Aminu B. Usman is the Associate Head of Computer and Data Science, York St John University in UK. He received a Computer Science degree from Bayero Universidad, Kano, a Master's degree in Network Security from Middlesex University, London, and a Ph.D. in Network Security from Auckland University of Technology. His current research is on Security and Authentication methods, IoT, next-generation networks and security issues in wireless networks.