DATA APPROACH TO BIOMETRICS IN CYBERSECURITY WITH RELATED RISKS

Angelo Arteche
Department of Computer Science and Engineering
Oakland University
Rochester, MI, USA
arteche@oakland.edu

Cody Asher
Department of Computer Science and Engineering
Oakland University
Rochester, MI, USA
casher@oakland.edu

Caroline Bull
Department of Computer Science and Engineering
Oakland University
Rochester, MI, USA
carolinebull@oakland.edu

Henry Dare
Department of Computer Science and Engineering
Oakland University
Rochester, MI, USA
hdare@oakland.edu

Isha Datey
Department of Computer Science and Engineering
Oakland University
Rochester, MI, USA
ishadatey@oakland.edu

Erich Elshoff
Department of Computer Science and Engineering
Oakland University
Rochester, MI, USA
erichelshoff@oakland.edu

Mohammed Mahmoud
Department of Computer Science
Bemidji State University
Bemidji, MN, USA
prof.mahmoud@bemidjistate.edu

***Abstract*—This paper will research the implications that biometrics have on cybersecurity. The paper will introduce topics such as what biometrics are, the recent advancements in it, and how biometric data is stored. The paper will then proceed to discuss privacy implications of biometric authentication, how it is implemented at an organizational level, its cost, and the performance impacts of using it. Subsequently, the paper will delve into what models and systems are currently used for biometric authentication and the risks posed by using it. The risks posed in paragraphs prior will be addressed in their own section, and the advantages will be presented to be weighed against the risks. The paper will also include discussions on how biometric data can be stored securely and processed using technology like cancellable biometrics. The paper will present the findings, and recap arguments brought forward throughout the paper.**

Keywords—biometrics, cybersecurity, privacy, biometric storage, biometric authentication, biometric authentication risks.

I. Introduction

Today's organizations are increasingly turning to biometric security in place of traditional authentication methods such as passwords and key cards in order to improve the secure protection of valuable data. However, without urgent advancements in accuracy and anti-spoofing techniques [23], biometric authentication runs the risk of offering institutions a false sense of security. For example, a 2017 study exploring the system limitations of fingerprint scanning, it was found that only 5 "master fingerprints" were necessary in order to unlock 65% of smartphone devices [1]. Other limitations which provide challenges to biometric authenticators include the threat of stolen biometric data, changing human features, and maintaining system accuracy across different populations. To illustrate the latter point, an infamous look into Amazon's facial recognition system found high levels of accuracy among fair-skinned users, which then declined rapidly when tested on users with darker skin tones. Google's voice recognition system notoriously recognizes male users at a 13% higher accuracy rate than female users [2].

In spite of biometric issues of inaccuracy and exploitability, they continue to be favored by organizations who've adopted them. Among reasons for their expanded use are convenience for the user (to no longer need to maintain a list of updated passwords), cost effectiveness (in prevented data breaches), as well as the uniqueness of common biometrics (although they may not be properly differentiated by a system, no two fingerprints are alike). Biometrics are also considered a necessity in light of the growing obsolescence of passwords, which are no longer considered an effective method for mitigating cyber-attacks regardless of their complexity [3]. The ongoing need for improved biometric authentication methods provides development teams with opportunities to engineer more secure, effective systems.

Two especially promising cutting-edge technologies in the field of biometrics can primarily be found in "voiceprints" (a spectrogram mapping of a voice uttering a specific word or phrase) and live facial mapping. Both promise a high level of accuracy, are increasingly difficult to spoof, and do not require the use of niche hardware to be installed on a device (in contrast to fingerprint or retina identification methods).

II. What Are Biometrics

Biometrics are the physiological and/or behavioral characteristics used for automatic recognition of an individual. Biometrics makes it possible to determine an individual's identity based on unique identifiers such as fingerprints, instead of the use of more traditional methods of identification, such as a driver license [4].

In order to be classified as a biometric, a characteristic must satisfy four requirements: Universality, Distinctiveness, Permanence, and Collectability [4].

979-8-3503-2028-2/22/$31.00 ©2022 IEEE
DOI 10.1109/CSCI58124.2022.00248

Universality requires the characteristic being considered to be present in the majority of the population that would be using the characteristic for identification. A characteristic with high universality would appear in everyone or almost everyone in a population, while a characteristic with low universality would appear in almost no one in a population. DNA would have high universality since it appears in everyone.

Distinctiveness requires a biometric to be sufficiently different between persons. Characteristics like weight and height would have low distinctiveness since there's a good possibility that multiple people would have similar height and/or weight. Fingerprints would have high Distinctiveness since it is very unlikely that two individuals would share a fingerprint.

Permanence requires a biometric to remain the same over a period of time. A biometric with high permanence would not change over time, while a biometric with low permanence would change frequently. Voice is an example of a biometric that could change very frequently depending on factors such as age, mood, and health.

Collectability is a measure of the ease in which a biometric can be acquired. A biometric with high collectability would be easy to acquire, while a biometric with low collectability would be difficult to acquire. Signatures are easy to collect and would therefore have high collectability [5].

For a biometric to be practical in the use of identification three issues should be considered: Performance, Acceptability, and Circumvention [4].

Performance considers the biometric potential accuracy and speed in identification, the resources required in order to reach a desired accuracy and speed, and any factors that would affect the accuracy and speed in the environment the biometric is being used.

Acceptability measures how willing a person is to present the biometric as a form of identification. A biometric with high acceptability means that most people will be willing to produce the biometric, while a biometric with low acceptability means that people would be unwilling to produce the biometric. The more a biometric intrudes on a person's daily life the less acceptable the biometric will be.

Circumvention represents how easily a biometric could be falsified in order to produce a false positive for the intended biometric. A biometric with high circumvention would be vulnerable to false positives while a biometric with low circumvention would be difficult to produce a false positive.

III. Recent Advancements in Biometrics

The field of biometric technology has been around for over 200 years, although it is only now slowly becoming more common and reasonable to use on everything from personal electronics, to major public spaces such as airports. Different biometric technologies have vastly different challenges, and needs on how accurate they need be with reasonable assumptions regarding their use case. The most recent, and readily available biometric technology to be consistently updated is what would be called “Machine-dependent” [6]. Most seen by Google, where their test involves tracking keystrokes and mouse movements as the end user selects images that match a query from a set. While that helps validate that users are not machines, and over time may build a database to help validate that certain movements could belong to a user, that doesn’t help validate them outside of technology.

Since the end of 2018, 17 airports across the United States now use facial recognition technology to help improve efficiency at airports [7]. Compared to the old method of manually checking visas and passports, this technology helped each user check in approximately 4 minutes faster. The technology uses photos of passengers walking up to enter the loading bay area, and compares them to photos available on the passenger’s driver’s license, or passport/visa. While the technology is quick and efficient, the designers say that for it to become a true and safe replacement of manual passport/visa checking, that the accuracy of the technology needs to be improved [7], and they should try harder to implement machine learning (ML) behavioral systems as well, as ensuring the machine doesn’t get to interact with the user directly.

Currently, when speaking of biometrics, the current most commonly used ones would be the user’s face, iris, fingerprint, and their voice [8].

Another currently worked on proposed model would be comparing user’s body parts to themselves. For example, comparing the left iris to the right iris of an individual, instead of just using one sample. Interestingly enough, in its current stage, the technology is unable to tell any more or less of a difference between two of the same individuals irises, and that of a stranger with a similar color. However, when an experiment was performed in person, humans themselves were 83% correct [8] in identifying whether or not images of an eye were from the same person, leading the creators to believe that what humans look at, is somehow not the same as what the technology is looking at.

Through testing of users’ vein patterns in their hands, it was found that vein patterns from the same user on their different hands are about 30% more similar than when comparing to a different individual [8], so they are now currently attempting to pour more data into the ML process to see if this is still a possible avenue for new biometric identifiers.

IV. Storage of Biometric Data

How is biometric data collected and stored, and how do we secure that data? When biometric data is first collected from the user, it should be noted that the information stored is not the literal image or recording of the user’s biometric, but rather a proprietary mapping of specific features such as the face or thumbprint. Along with providing an efficient means of encoding the biometric into machine-readable data, this provides a degree of privacy to the user as their credentials are not directly associated with their identity. This mapping is known as the biometric template [9]. The biometric template must be continuously updated with new samples of validated

data in order for the user to continue being authenticated as their features gradually change. The storage of biometric data can be broken down into two stages, enrollment and verification. The enrollment stage is the physical capturing of the user's biometric information when they attempt to gain access. The verification stage is the process of updating the biometric template with a new validated sample, provided the new sample matches the template enough to be considered genuine [10].

Although biometric authentication has been developed as a tool for securing sensitive data, the secure storage of biometric data is a field of defensive cybersecurity unto itself. Protecting biometric data from potentially malicious actors falls within two categories, its physical storage location as well as the ways in which access of the data can be obfuscated, such as encryption.

Possible locations for the storing of the biometric template can be selected from three options: on a user's local device, via a remote server, or distributed within a cloud architecture. Of these, the most difficult location for an attacker to gain access to would be on the local device, particularly in a chip kept isolated from any network activity as is the case on many smartphones. Although it is by far the simplest method, its value should not be understated, as the only way to get ahold of the user's biometric information would be through the seizure of the device [11]. A disadvantage of this storage method would be the constraint it imposes on the developers of the biometric system, as they would be unable to utilize user data for analysis. This would then restrict resources of raw data available when iterating upon the algorithms which comprise the biometric template. By contrast, the most exploitable location is on a remote central server. Due to its vulnerability as a single point of failure, remote centralized servers are gradually being abandoned by organizations in favor of distributed cloud storage, which merges the advantage of data being difficult to access at a single source with the convenience of having data connected to a network for backup and analysis. Distributing biometric data across the cloud also provides additional privacy protection for end-users, removing any direct association between their identity and their stored biometric template [12]. In short, the client versus server model for storage of the biometric template forces a crucial decision for organizations considering biometric authentication. The distributed server model provides the clear advantage to companies improving their system, whereas a locally stored template offers the highest level of security and privacy of end user personal data to client organizations.

V. Biometrics in Authentication

Biometrics are becoming substantially more prevalent in authentication scenarios. The benefits of using biometric data ensure that passwords are neither lost nor forgotten nor expire. The most popular methods of biometric authentication involve using facial recognition, voice authentication, fingerprint scanning, and iris scanning [13]. Using any of the above methods make it substantially harder for a potential attacker to gain access to a victim's information.

Specific biometrics including facial scans, voice identification, fingerprint scanning, and iris scanning are the most prevalent authentication options. One advantage of using biometric data for authentication is the substantially decreased risk of successful hacking attempts. Organizations that choose to implement biometric authentication have the advantage of using it in combination with a standard password, or having the password as a backup. This combination allows for effective multi-factor authentication which increases system security even further.

Historically, a weak point within systems and security has always been users, with their tendency to fall for phishing attacks as well as using weak, easy to guess passwords. Implementing biometric authentication helps to reduce, if not eliminate, this risk since biometric data is much harder to replicate [14]. In combination with the secure storage of this data, it even makes phishing attempts far less useful, as the received data is no longer a character string representing a password, but instead a large collection of information which combined, represents the biometric identity. If a hacker is not prepared or able to decode this information, the received password becomes unusable and stops a potential attack despite a user essentially providing their password.

A disadvantage of biometric authentication is the integration cost. In order to implement biometric authentication, an organization must purchase and install the hardware capable of reading or scanning the biometric data. This hardware must be installed at any workstation where the user will authenticate with biometrics. Beyond the hardware, the backend will need to be set up to both prompt for authentication based on the organization's policy, e.g. once a day or once a month, then there must be a server and software capable of reading, matching, and approving the received biometric data from the user. All of this infrastructure is a reason biometric authentication is not as widely implemented, it's far safer but comes at a high installation cost.

With many biometric authentication methods requiring additional infrastructure and end user hardware, organizations may find the voice authentication to be the most appealing, since almost all modern computers have built in microphones which eliminates the need for additional hardware. This method of authentication is also one of the most thoroughly tested, utilizing ML and Python algorithms to match a user's voice and eliminate spoofing attempts.

A study was conducted on one such open source authentication method, Mozilla Common Voice. This authentication method is built upon the Python language and was tested against features unique to a user's voice, in this case Mel frequency Cepstral Coefficients (MFCC). Using voice data samples available on the Carnegie Mallon University website, the Common Voice authentication model was able to achieve up to 89.20% authentication accuracy [15]. This study shows that organizations may adopt less expensive, open source options, to receive the security benefits of biometric authentication without having as great of an implementation cost.

A final important consideration when implementing biometric authentication is the performance degradation on

servers and end user devices. Since biometric data is not as simple as matching an entered password to a stored database entry, there needs to be specialized software and algorithms configured on an organization's servers. The end users will then need programs capable of interacting with the biometric data and the authentication server. Depending on the efficiency and level of security, the authentication software can range from a negligible performance impact up to a noticeable performance degradation [16].

VI. Risks in Biometrics

Despite the growing interest and rapid growth in biometrics as well as the immense predicted future potential, the technology does not come without its own set of risks. Using biometric data to access our personal devices is increasing as a way to get around the limitations of the commonly used password-based mechanism: it is easier, more convenient, and (theoretically) more secure. But biometric data can also be stolen and used in malicious ways.

One such prevalent example is fingerprint biometric authentication, used increasingly in modern smartphones for its size and affordability. It is simple and quick, however, it is vulnerable to attacks as hackers can steal fingerprint biometric data [17]. Stealing and using fingerprint data is not as easy as lifting someone's social security number but experience and history tells us that once something is used extensively, criminals will figure out how to misuse and monetize it.

Most proposed techniques to resolve this risk of hackability, from a quantification, storage, and communication point of view, are designed for discrete data and use simple similarity measures. However, true biometric data requires complex similarity functions. Also, the techniques designed for real-world biometric data are either ad hoc and without formal proof of security or don't provide a sufficiently rigorous security formulation [18].

Focusing on fingerprint authentication, the fingerprint is not only used to unlock mobile devices but also to access many banking applications that are progressively implementing fingerprint authentication for payments. Since a fingerprint, like most other biometric factors such as the human iris and face, lasts for life. Once compromised, it could not be recovered or changed like traditional password recovery mechanisms, where the victim can replace a stolen password with a new one. Since the fingerprint is also associated with many other identifications, therefore, the biggest concern of users is how secure the fingerprint architecture is designed [17].

Voice authentication is also another popular example of biometric authentication currently in use. The proposed model above focuses on a voice authentication model. But the model also presents a few problems. Most importantly, as addressed in the paper, the software module produced high error values when the user speaks with a very different intonation and/or volume than when registering. For example, the system may not recognize it if the training sample is very small and monotonous. The dataset plays a very important role in deciding how accurate the voice detection will be. This is not a particularly dangerous risk for the data stored on the device, but it may cause inconvenience for a legitimate user. One of the solutions suggested is to set up the special architecture of authentication service. For example, after a certain time, you can ask the device user to supplement the training sample. Another problem is that the system has not been tested in real conditions. It will be interesting to see how it will be implemented in real conditions in the future [15].

However, ML solutions can be vulnerable to adversarial attacks. Feature reduction is one of the fundamental tasks in ML aimed at controlling overfitting. However, in adversarial tasks, feature reduction can allow the adversary to evade the classification system. This model uses an open source dataset with a controlled selection of features for extraction for the ML procedure. It may allow adversaries to evade the model [19].

Human error is another one of the most critical risk factors in the security of any piece of technology [20]. By simply uploading their picture to Facebook or using their thumb to unlock their smartphone, users may be giving away critical data without realizing where the information is going and what it is being used for [18]. Several studies have described that the end user fails to understand the permission warnings that malicious software wants to have during installation. This ultimately enables an attacker to gain super user system level privileges to use system resources remotely, such as the camera of a device [17].

Another example of human error are phishing attacks, which are dangerous to biometric applications because the end users might be tricked to download malicious software that looks safe to download. These phishing applications authorize imposters with super user rights to use the device. Some of the phishing applications might use the user's facial photographs or stored fingerprints for unacceptable purposes, which may raise great privacy concerns regarding the use of these applications.

Privacy issues are at the heart of the ethical issues of Biometrics. In 2001, the US RAND Corporation discussed two privacy issues in the application of biometrics in their report: information privacy and physical privacy. Biometric information is collected through observations of individuals and is used to identify individuals by things such as fingerprints, faces, hand shapes, DNA, etc., which are undoubtedly personal information. However, it is controversial whether the biometric information stored in the biometric system is still personal information [21].

Currently there is no restriction on what biometric information companies can share and with whom. As technology advances, we will encounter privacy and security issues even more frequently. It is within reach for companies to use new technology to replace all passwords, security personal identification numbers, access codes, etc. But these advances might allow companies to "go too far" with a person's biometric data, giving unprecedented access [18]. Companies are not the only institutions at risk of breaching someone's privacy. It is feasible to envision a society where the government has extensive biometric data, most specifically facial recognition data, that can be used for the purpose of illegal spying [18].

Another important ethical issue is autonomy, and an important part of exercising autonomy is informed consent. In order to ensure the individual's informed consent, it is important that the individual understands the purpose and meaning of the biometric system. In general, adults are considered to have sufficient ability to understand information. The problem is mainly surrounding the child's informed consent when using biometrics. Similar informed consent issues also come from vulnerable populations, such as the elderly, minors, or those with certain medical conditions.

On top of that, social exclusion is also an important ethical issue to be considered. At least for now, biometric acquisition devices are not capable of handling individuals other than normal values, and some individuals are not able to be identified and thus excluded. Especially when these systems are linked to social welfare, these unidentifiable individuals are likely to be excluded from social welfare, leading to injustice. These groups include: people with disabilities or educational limitations, the elderly, people of certain races, and homeless people.

Moreover, applications having biometric identification systems involved are prone to several types of attacks, which are classified into two main categories: direct and indirect attacks [17].

Direct attacks often involve spoofing and alteration while indirect attacks include phishing attacks, malware injection, and circumvention. Spoofing attacks involve gaining an invalid and unauthorized access to biometric mobile applications by presenting a fake biometric feature or trait such as face mask, silicon finger etc. Spoofing attacks are the most popularly employed attacks to breach information.

Alteration attacks refer to presenting the imposters' own information with alterations using various falsification techniques such as obliteration, imitation, and distortion. One instance of facial features alteration is applied by using facial plastic surgery. Related to this, an attacker can imitate facial signatures of any politician by taking his pictures from various angles. Other studies have discussed scanner limitations and high profile attacks like artificial fingerprints made by different materials having similar characteristics as human skin, which have been deceptively accepted by many scanners as well [17].

Indirect attacks can be initiated in two ways: on the software modules or on the interface between modules. Interface between modules attack is also named as replay attack. In this kind of attack, an attacker attempts to resubmit the stolen information intercepted previously before or after applying extraction. An example of such an attack is stealing the transmission information from the transmission channel between a comparator and database to compromise the results. While referring to attacks on software modules, a malware injection attempt can be made between feature extractor module and comparator module to produce desired outcomes [17].

VII. Addressing Biometric Risks

Using biometrics as a tool for cyber security is an intriguing prospect for many organizations looking to advance their systems. Biometrics can be very beneficial, for example allowing for more convenient and efficient authentication, however, they also come with several disadvantages and risks that must be managed. In this section we will discuss the potential solutions to these challenges and risk mitigation strategies related to biometrics.

The primary risk factor for a biometric system is an attack from a malicious third party. These attacks can be devastating and mitigating the risks should be a primary concern when building a system. As previously discussed, there are two types of attacks related to biometric systems: direct and indirect. Direct attacks target biometric technology as the entry point. Indirect attacks gain access to the system through other means.

Preventing indirect attacks is done by developing a comprehensive security system. Since biometrics is a tool that can be used within a security system, rather than a comprehensive security system by itself, it is important to note that the deployment of biometric technology and the security implications will often be unique to each system. Risk mitigation strategies should begin at the highest level, looking at the entire system and analyzing the potential threats, vulnerabilities, asset priorities, and finally the mitigation techniques. By understanding these components of a security program, you can apply elements of protection, detection, and reaction to the system. Integrating biometrics into a system is a way to bolster the protection portion of a system, but there are implications in the other portions of the system. For example, if a biometric fingerprint scanner is added as an authentication measure, not only does this bring the security risks of the scanner itself, but considerations for the entire system should be made. This includes having proper protocols in place and trained staff to uphold the policies and maintain the system. When implementing biometrics into a system it is crucial to create a robust and holistic security system that can complement the biometric technology.

While creating a comprehensive security system is fundamental to mitigating risk, there are also advancements in specialized countermeasures specifically related to biometrics. These countermeasures try to prevent direct attacks, which target the biometric technology as the weak point to gain entry to the system. According to research, without any countermeasures in place, all biometric systems are vulnerable to spoofing [22]. Further, creating and implementing countermeasures for spoof attacks has proven to be difficult.

One technique to counter spoofing is known as liveness detection. Liveness detection attempts to determine if the biometric data given is from a genuine human or an imposter through either hardware or software additions. Liveness detection through hardware involves implementing additional sensors that can add data points to biometric scans for features like perspiration or blood pressure. Scanning for additional bio-markers makes spoofing significantly more difficult but will increase the cost, complexity, and processing power

required. Liveness can also be detected through software that analyzes biometric samples to determine authenticity [26].

One promising approach to software-based liveness detection is using convolutional neural networks (CNN), a deep learning (DL) technology trained to compare images. In a study done in 2019, AlexNet, a convolutional neural network architecture, was used to compare images of authentic and fake fingerprints. The AI's ability to distinguish real and spoofed biometrics was compared to conventional methods that use hand-crafted data markers . This study found that using DL methods was highly efficient, outperforming conventional methods. The use of CNN consistently showed to have lower error rates than traditional methods across a number of different biometric sensors [24].

Liveness can also be detected using a technique called challenge-response authentication. This technique is also used to authenticate non-biometric data, often asking users to provide additional information (e.g. mother's maiden name) which is verified against the database. This technique can be used with biometrics as well. Users will be asked to perform a specific action at the biometric scanner, like blinking or turning their face. It is difficult for an attacker to successfully produce the required response, making this an effective technique to stop spoofing attacks. While this technique can increase the security of the system, this solution is invasive to the user, can be time consuming, and has the potential for user error [25].

Another risk biometrics creates is the potential for attackers to gain access to personal biometric data. Since biometric data is permanently linked to a person and can't be changed or reset like a password, this creates a great privacy risk. To combat this risk, a strategy called cancellable biometrics has been developed. Cancellable biometrics refers to the distortion of biometric data in such a way that attackers are unable to recognize the data but can still be used for authentication purposes. This is a type of visual cryptology. Cancellable biometrics works by first taking a biometric scan of the user to create a template. Before the template is stored, the system will extract the specified features from the image, then the image is warped using a technique like hashing and stored in this distorted state. The authentication process works by repeating this process at the scanner and authorizing the user if the templates match. This allows the biometric data to be canceled in case the data is compromised in an attack. Cancellable biometrics is still an emerging technology but solves a major privacy concern[26].

Both cancellable biometrics and software-based liveness detection are promising technologies that could help improve the security and privacy regarding biometrics. However, to be viable, they should also be able to be deployed in a cost-efficient manner. These biometric systems have high computational costs and require fast response time with little to no errors, which means there will always be significant hardware costs to support these requirements. These costs include computational hardware and storage. Recent advancements have shown that moving these technologies to a cloud platform is more cost efficient. A recent study used Amazon Web Services (AWS) cloud servers to host their cancellable biometric systems. Their novel system used the cloud to run the bulk of computations including the deep learning module, the database, and the biometric engine. The total cost of the system was around $20,000 per year running non-stop. This is significantly cheaper than established biometric solutions. Further, the cloud service allows for scalability. The study found that moving to the cloud not only saved money, but also showed higher performance compared to a stand-alone system [27].

Overall, there are many ways to prevent attacks and reduce risks in biometric systems. Ideally, some combination of liveness detection through hardware and software with cancellable templates will be the most secure. However, this can be both costly and complex. The needs of every system will be different, so it is important to run a cost-benefit analysis on the best security solutions for each system.

VIII. Results

To efficiently summarize the proposed models section, we defined biometrics as "the automatic recognition of individuals based on their physiological and/or behavioral characteristics." Biometrics in principle have been around for over 200 years, but have recently been increasingly incorporated into hardware and software security in the public sector. Basic fingerprint tests are now implemented in more and more personal devices every day, and while they do a decent job for the most part, there is still a decent risk of your authentication methods being spoofed, or your own biometric data stolen.

The storage of biometric data itself and the ability to protect that data from malicious actors is essential to biometrics. Biometric data is typically stored in three locations: the user's local device, a remote server, or within a cloud architecture. The most secure of these locations being the user's local device. When storing biometric data it is imperative to not store the literal of the characteristic being used, but rather to store the identifying features which then must be encrypted.

Biometric data is not just something that we can forget about, and it is not something that we can perfect and then throw away. Although it still has issues, and is still in its' early stages of development, biometric authentication is all around our personal devices, and soon with extra security, it will likely be all around corporate lives as well. While you may be able to forget a password, aside from extremely low percentage accidents, you will always have your biometric data; which makes it a very simple solution to use it for personal device authentication where you currently don't need as much security as a big corporation would, as their data is much more important, and desired.

IX. Conclusion

In spite of biometric authentication's increased prevalence, several key weaknesses make its universal adoption problematic for organizations. These reasons include upfront cost of implementation, potential leaking of sensitive user data, and the inaccuracy of biometric authentication algorithms in identifying legitimate users, as well as their susceptibility to the spoofing of credentials.

These are all prominent reasons why biometric authentication is not the standard. Biometric authentication can also be considered a risk as they can become a single point of failure for multiple accounts security. However, for the same reason they remain increasingly popular for users, due to the convenience of adopting them versus the inconvenience of maintaining a collection of passwords. In correlation, as research and support for biometric authentication grows, it becomes more available to smaller organizations which causes a snowball effect, encouraging the methodology to be further developed and implemented.

Biometric authentication poses certain risks and restrictions, but the security advantages it provides offers great incentive for organizational and personal implementation. Advantages such as eliminating weak passwords, and changing the way passwords are stored and validated, greatly assist in minimizing security vulnerabilities caused by negligent users. The parallel development of anti-spoofing systems provides some protection against intrusion attempts.

X. Future Work

There is still a lot of work to be done in this booming sector of biometric authentication. This rapid technological advancement needs to be balanced with cybersecurity in the coming years.

For the future, we hope to explore in our research more proposed solutions and models based in the Python scripting language. Currently, we lack the technical knowledge to work on and propose a model from scratch, but there is a wide future scope to research, study and devise a new model. The model proposed by Sidorova et al. is a great, informational basis to build any future models on. Considering that some of our researchers have some background in ML using matlab, it was a model we could comprehend more than others.

In addition to that, the field of AI in biometric authentication has an extensive scope for improvement and questioning in the discipline of cybersecurity as one of the most important and crucial concerns about replacing text-based authentication with biometric authentication is security. It is not massively efficient as of yet but ML and automated AI systems or self-learning AI systems are excellent areas to focus on to achieve exemplary results.

Besides that, it is very important to address the ethical issues mentioned in the paper, such as privacy issues and consent issues etc. and find solutions to combat them. Along with that, social inclusion is something to be worked on. One of the possible solutions may be to introduce more extensive datasets with equal chunks of data dedicated to each kind of group to train ML models for biometric detection.

We would also like to work on this paper more extensively given more time and cover more presented solutions and models to take care of the risks listed above, as well as to improve the implementation and storage of biometrics in authentication in this literature review in the future.

References

[1] A. Roy, N. Memon and A. Ross, "MasterPrint: Exploring the Vulnerability of Partial Fingerprint-Based Authentication Systems," in IEEE Transactions on Information Forensics and Security, vol. 12, no. 9, pp. 2013-2025, Sept. 2017, doi: 10.1109/TIFS.2017.2691658.

[2] Tatman, Rachael. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. 53-59. 10.18653/v1/W17-1606.

[3] Routh, Jim "The Growing Obsolescence of Passwords"

[4] A. K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4-20, Jan. 2004, doi: 10.1109/TCSVT.2003.818349.

[5] Andress, J. (2014). The Basics of Information Security: Understanding the Fundamentals of InfoSec in Theory and Practice (2nd ed.). Syngress.

[6] Tumpa, Sanjida Nasreen,et al. "Social Behavioral Biometrics in Smart Societies." Advancements in Computer Vision Applications in Intelligent Systems and Multimedia Technologies, edited by Muhammad Sarfraz, IGI Global, 2020, pp. 1-24.
https://doi.org/10.4018/978-1-7998-4444-0.ch001

[7] Nimra Khan and Marina Efthymiou. "The use of biometric technology at airports: The case of customs and border protection (CBP)." „ DCU Business School, 6 Nov. 2021, https://www.sciencedirect.com/science/article/pii/S2667096821000422

[8] KUZU, R. S.; MAIORANA, E.; CAMPISI, P. On the intra-subject similarity of hand vein patterns in biometric recognition. Expert Systems with Applications, [s. l.], v. 192, p. N.PAG, 2022.

[9] Biometric Encryption, http://www.cse.lehigh.edu/prr/Biometrics/Archive/Papers/BiometricEncryption.pdf

[10] how is biometric data stored? https://www.nec.co.nz/market-leadership/publications-media/how-is-biometric-data-stored/

[11] "The Standard for Biometric Data Protection"Journal of Law & Cyber Warfare Vol. 7, No. 1 (FALL 2018), pp. 61-84 (24 pages)

[12] Julien Bringer and Hervé Chabanne. 2010. Negative databases for biometric data. In Proceedings of the 12th ACM workshop on Multimedia and security (MM&Sec '10). Association for Computing Machinery, New York, NY, USA, 55–62.

[13] Bhattacharyya, Debnath, et al. "Biometric authentication: A review." International Journal of u-and e-Service, Science and Technology 2.3 (2009): 13-28.

[14] Jain, Anil K., and Karthik Nandakumar. "Biometric authentication: System security and user privacy." Computer 45.11 (2012): 87-92.

[15] A. Sidorova and K. Kogos, "Voice authentication based on the Russian-language dataset, MFCC method and the anomaly detection algorithm," 2020 15th Conference on Computer Science and Information Systems (FedCSIS), 2020, pp. 537-540, doi: 10.15439/2020F43.

[16] Sarkar, Arpita, and Binod K. Singh. "A review on performance, security and various biometric template protection schemes for biometric authentication systems." Multimedia Tools and Applications 79.37 (2020): 27721-27776.

[17] M. R. Zafar and M. Ali Shah, "Fingerprint authentication and security risks in smart devices," 2016 22nd International Conference on Automation and Computing (ICAC), 2016, pp. 548-553, doi: 10.1109/IConAC.2016.7604977.

[18] Cooper, Isaac and Yon, Jimmy, Ethical Issues in Biometrics (September 11, 2019). Sci Insigt. 2019; 30(2):63-69, Available at SSRN: https://ssrn.com/abstract=3451985

[19] https://proceedings.neurips.cc/paper/2014/file/8597a6cfa74defcbde3047c891d78f90-Paper.pdf

[20] Hughes-Lartey, K., Li, M., Botchey, F. E., & Qin, Z. (2021). Human factor, a critical weak point in the information security of an organization's Internet of things. Heliyon, 7(3), e06522. https://doi.org/10.1016/j.heliyon.2021.e06522

[21] N. Memon, "How Biometric Authentication Poses New Challenges to Our Security and Privacy [In the Spotlight]," in IEEE Signal Processing Magazine, vol. 34, no. 4, pp. 196-194, July 2017, doi:10.1109/MSP.2017.2697179.

[22] Rhodes, K. A. (2003). Information security: Challenges in using biometrics. General Accounting Office.

[23] Hadid, A., Evans, N., Marcel, S., & Fierrez, J. (2015). Biometrics systems under spoofing attack: an evaluation methodology and lessons learned. IEEE Signal Processing Magazine, 32(5), 20-30.

[24] Marcel, S., Nixon, M. S., Fierrez, J., & Evans, N. (2019). Handbook of Biometric Anti-Spoofing. Springer.

[25] Khan, M. K., Zhang, J., & Alghathbar, K. (2011). Challenge-response-based biometric image scrambling for secure personal identification. Future Generation Computer Systems, 27(4), 411-418.

[26] Kumar, N. (2020). Cancelable biometrics: A comprehensive survey. Artificial Intelligence Review, 53(5), 3403-3446.

[27] T. Sudhakar and M. Gavrilova, "Cancelable Biometrics Using Deep Learning as a Cloud Service," in IEEE Access, vol. 8, pp. 112932-112943, 2020, doi: 10.1109/ACCESS.2020.3003869.