Development Project Research

Privacy-Preserving Biometric Authentication

P2665421 - Luke Dawson

CTEC3451

Contents Page

Introduction

In an era where data security and personal privacy are at the forefront of global discourse, the need for robust and secure authentication applications has never been more impactful. Biometric Authentication, leveraging innate human characteristics like fingerprints, iris patterns, and voice signatures, offers an intuitively secure solution. However, its strength in using unique, unchangeable biological traits, is also its Achilles' heel: unlike traditional passwords, once compromised biometric data cannot be reset. This vulnerability is what makes the protection of biometric data paramount in modern society. This is where the introduction of Privacy-Preserving Biometric Authentication (PPBA) is key, PPBA is an innovative approach that seeks to combine the efficiency of biometric systems with the rigorous security of cryptographic techniques, such as homomorphic encryption. This research delves into the intricacies of PPBA, exploring its foundational concepts, its integration with various encryption schemes, and its potential to revolutionise the future of secure authentication.

Legal and Ethical Implications

In the evolving landscape of biometric technology, legal and ethical considerations have emerged as crucial concerns. Historically, the development of biometric systems primarily operated in a bubble, devoid of specific legal oversight. However, as society became more data-driven the urgency for legislation became more pronounced. An examination of the initial legal discussions reveals a tapestry of debates oscillating between the promise of heightened security and concerns over individual privacy (Drozd, A., & Pieprzyk, J. (2017)). “Biometric security through the lens of its legal and ethical underpinnings” *International Journal of Information Management, 37(6),(601-609).* Drozd and Pieprzk’s analysis underscores the balancing act regulators face ensuring the potential of biometrics is harnessed while safeguarding against potential misuse.

Throughout recent years biometric systems have shifted to have a bigger emphasis on transparency when it comes to user content. Due to technology advancing at unprecedented rates, stakeholders are having to become acutely aware of the necessity for clear communication about data use and storage protocols ***(Jain, A. K., & Ross, A. (2016). "Bridging the Gap: From Biometrics to Forensics". Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1705)).*** Jain and Ross highlight the imperative nature of transparency, stating that in the absence of clarity, the trust needed for these systems can quickly disperse, rendering them ineffective and counterproductive.

Lastly, the universality of biometric systems mandates a standardised approach. Standardisation not only streamlines system deployment but also fortifies their security framework. With the individuality that biometric data brings, standardised protocols play a crucial role in ensuring consistent performance and reliability across diverse environments ***(Sutcu, Y., Sencar, H. T., & Memon, N. (2007). "A Secure Biometric Authentication Scheme Based on Robust Hashing". Seventh International Workshop on Multimedia Signal Processing, 70-73).***

Biometric Authentication

Brief History of Biometric Authentication

Biometric authentication is a relatively new security protocol only being used as tools for identifying and security *"...in the late 19th century with the work of Alphonse Bertillon"*[1] biometrics themselves, however, can be traced back to ancient times when fingerprints and handprints were used as signatures and seals. Voice, iris and facial biometric systems were only recently developed being implemented in the *"latter half of the 20th century"*[1]

Types of Biometric Authentication

According to the *"Biometrics Institute"*[2] there are 16 different types of biometrics; DNA, Ears, Iris, Retina, Scleral Vein, Face, Finger geometry, Fringer / Palm print, Gait, Hand geometry, Heartbeat, Keystrokes. Odour, Signatures, Vascular, and Voice.

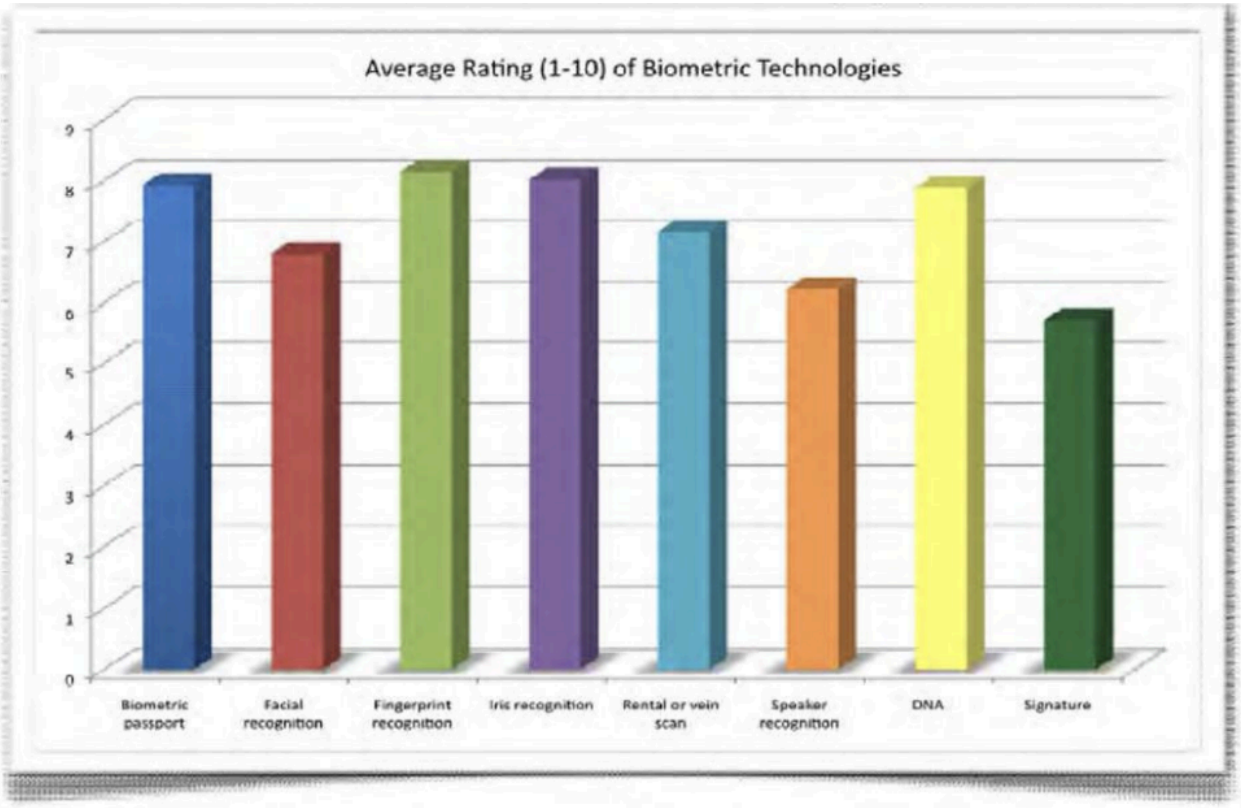


Figure 1. Shows the choosing methods of ordinary people when discussing biometrics (Figure from Biometric Authentication. Types of biometric identifiers)[5]

Benefits and disadvantages to Biometrics

The greatest appeal and advantage to using biometrics for authentication is the increased security biometrics provide, while the second main selling point is the accuracy provided with using biometrics for authentication. However, these strengths come at the high implementation costs, and invasion of privacy to its users.

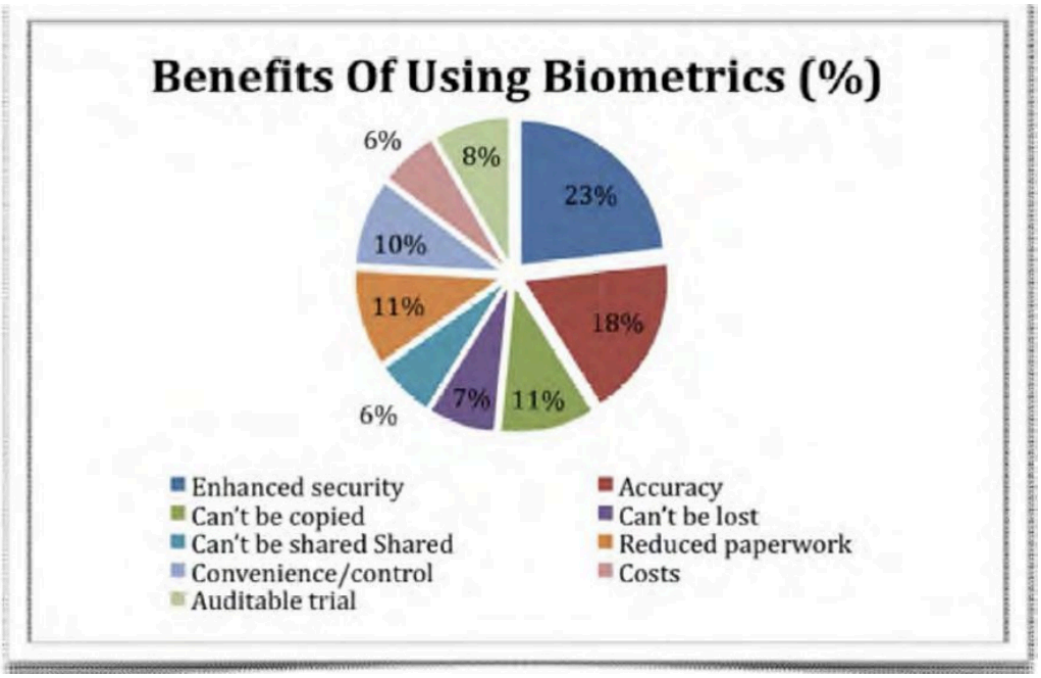


Figure 2. Strengths of Avantages (Figure from Biometric Authentication. Types of biometric identifiers)[5]

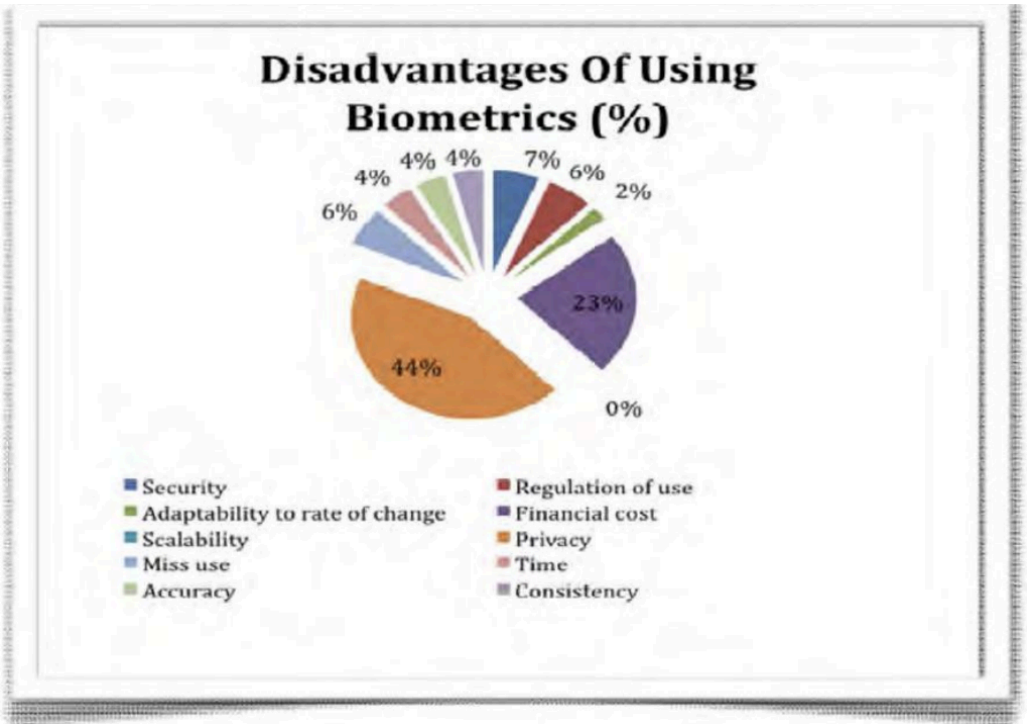


Figure 3. Impact of disadvantages (Figure from Biometric Authentication. Types of biometric identifiers)[5]

Along with the high implementation cost the technology that is currently available, builds challenges for the practicality of a

majority of the biometrics mentioned above. As stated in "*Biometric Authentication: A Review*"[3] *(Debnath, Rahul et al), "…thorough research it can be concluded that approaches made for simultaneous authentication and verification is most promising… But whenever the method we choose, main constraint will be its performance in real life situation…".*

Functionalities of Biometric Systems

There are three main steps when operating a biometric system these are; Verification, Identification and Screening.

Verification ("Is this person who they claim to be?"): Biometrics focuses on checking the authenticity of individuals through the input of biometric samples and data. For example, ***"a person claims that he or she is known as John Doe within the authentication system and offers his or her fingerprint; the system then either accepts or rejects the claim based on a comparison performed between the offered pattern and the enrolled pattern associated with the claimed identity.[6]".*** Examples of verification processes include: computer network logons, ATMs, and medical records management.

Identification ("Is this person in the database?"): Once the biometric input has been given it is then matched against the sample stored in the database, the identification then determines if the input and the stored sample match. Examples of the identification process include: National ID cards, Border control, and criminal investigation.

Screening ("Is this person malicious?"): Screening applications determine whether a person belongs to a watchlist of identities. Although the screening process is the last step in the biometric functionality life-cycle, it is the step with the most faults and potential errors. ***"By their very nature, the screening applications: 1) do not have well-defined "User" enrollment phase; 2) cna expect only minimal control over their subjects and imaging conditions; 3) require large sustainable throughput with as little human supervision as possible. Screening cannot be accomplished without biometrics[6]".***

Biometric Data in PPBA

Acquisition of Biometric Data

The gathering of biometric data is not yet perfect in practical situations, this is because each and every time the sample is acquired there are slight deviances that cause variations in the biometric signal. For example, nonuniform contact creates poor quality in the results of the data. This is because the ridge structure of a finger can only be captured entirely if the ridges that belong to the finger being imaged are in complete contact with the image acquisition surface and the valleys don't make any contact with the surface. However, "...dryness of the skin, shallow/worn-out ridges, skin disease, swear, dirt, and humidity in the air all confound the situation..." These outside factors will result in a nonideal contact (Figure 4).

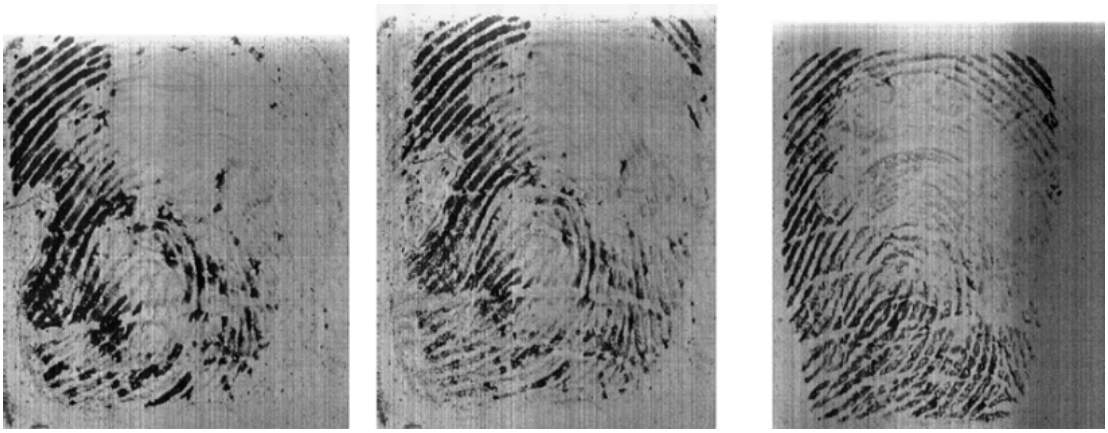


*Figure 4. Imperfect and unideal acquisition of the sample due to **"poor quality ridges"** and **"extreme dryness"**. (Figure from Biometrics: a tool for information security).*

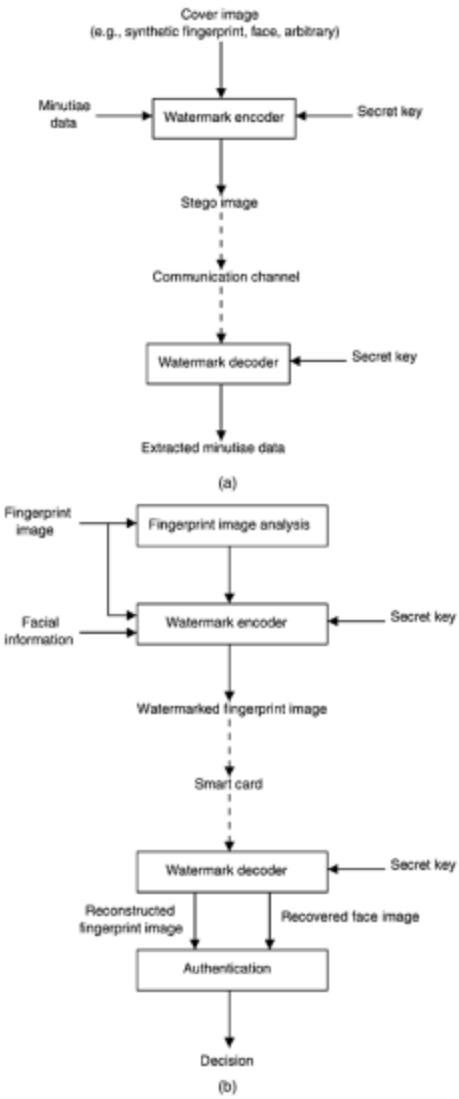
Similar to having dry, dirty or poor quality ridges, poor distribution or lack of inking of the finger will often result in "noisy" low contrast images. These errors can later lead to ***"spurious or missing fingerprint features[6]".***

The effectiveness of PPBA systems relies significantly on the quality of the biometric data it processes. Upon acquiring biometric data, an immediate assessment of its quality becomes imperative. As established by the ISO/IEC standards on the biometric performance testing and reporting, stringent criteria determine whether a captured biometric sample is of accepted quality. Compromised or tainted samples may lead to inaccuracies in the authentication process or even potential breaches, thus their

detection and possible rejection or re-capture are mandatory. To get the data ready for storage or matching we need to use an application called pre-processing. Bolle, Connell, & Ratha have thrown light on the relationships between various biometric accuracy metrics, emphasising the role of techniques such as feature extraction and normalisation in enhancing the reliability of biometric systems. These steps serve as foundational pillars, ensuring that the data entering the system is both authentic and optimised for operations, whether that be storage, matching, or encryption in the PPBA scheme.

Storage of Biometric Data

In the realm of Privacy-Preserving Biometric Authentication (PPBA) systems, the storage of biometric data is of paramount importance. Because of this, unlike traditional biometric storing techniques, PPBA systems follow a mathematical representation when storing the data. This is known as a template, as elaborated by Jain & Uludag in their pivotal work on hiding biometric data.



Once this template is generated, ensuring its privacy is of utmost importance. This issue is solved with the introduction of homomorphic encryption (HE), a technique highlighted by Gentry's groundbreaking research, which offers a solution, allowing operations to be conducted directly on encrypted data without needing the decryption process *(Further explained under the Homomorphic Encryption section below)*.

Jain & Nandakumar emphasise the intricacies of the system security and use privacy in biometric authentication, highlighting the importance of efficient retrieval and matching the extensive biometric databases. Given the vulnerabilities associated with centralised databases, Rathgeb & Uhl explore decentralisation, contemplating the use of innovative technologies evolving technological landscapes, continuous authentication has emerged as a topic of interest.

Figure 5. Diagrams of application scenarios (Figure from Hiding Biometric data[7])

Privacy-Preserving Techniques

What is Homomorphic Encryption (HE)?

Homomorphic Encryption (HE) facilitates the execution of arithmetic operations, including addition and multiplication, on encrypted data, removing the need for decryption to its plain text form. After performing these operations and subsequent decryption, the result aligns precisely with the outcome that would have been achieved had the operations been executed on the decrypted data directly. THis capability ensures that operations on the encrypted data do not compromise the confidentiality of the original plain text (Gentry, 2009).

HE stands as a key pillar in the edifice within the data privacy bubble. For example, in scenarios where data is being processed by third-party entities, yet granting them direct access to raw, unencrypted data remains undesirable. If you implement HE, data proprietors can provide these entities with an encrypted version of your data enabling them to perform necessary operations without ever gaining insights into the raw data.

This method is especially advantageous for organisations that leverage cloud computing, where the idea of ensuring data confidentiality becomes paramount. Through HE, data can be processed on cloud platforms, mitigating concerns surrounding potential exposure of sensitive information (Gentry, 2009).

Classification of Homomorphic Encryption

HE can be split into one of three distinct categories:

Partially Homomorphic Encryption (PHE): PHE supports singular operations, either addition or multiplication, but not both concurrently. An example of PHE is the ElGamal encryption scheme (Elgamal, 1985).

Somewhat Homomorphic Encryption (SWHE): SWHE permits a finite set of both additions and multiplications. However, upon exceeding a predetermined number of operations, an element termed as ‘noise’ is introduced. Noise causes the data to undergo a “refreshing” process.

Fully Homomorphic Encryption (FHE): FHE allows for boundless additions and multiplications on encrypted data. This concept, once deemed theoretical, was actualised

in 2009 through the pioneering work of Craig Gentry and the introduction of a process called bootstrapping.

Challenges and solutions of Homomorphic Encryption

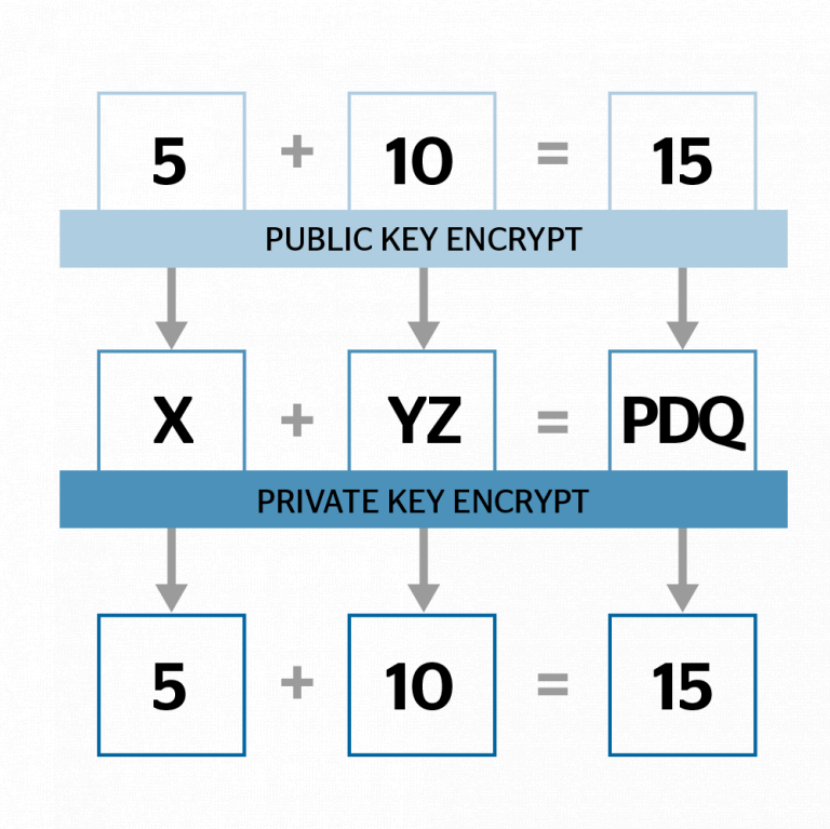


Figure 6. Example of how Homomorphic Encryption functions[15]

HE presents promising solutions to data privacy concerns. However, its implementation is not without challenges. Several issues arise, including computational overhead, increased storage demands, and the intricate management of noise accumulation.

The foremost challenge is the computational overhead. The operations on encrypted data, inherent to HE, are considerably more computationally demanding compared to operations on plaintext data. This intensity can lead to processing times that are multiple orders of magnitude slower compared to standard. This could cause constraints when trying to use HE in real-time or extensive situations. To counter this issue the use of optimised algorithms are used. These algorithms are meticulously designed to streamline and expedite the computation on encrypted data and can significantly reduce the computational cost. By minimising redundant operations, utilising efficient data structures, and strategically leveraging mathematical shortcuts, optimised algorithms can expedite the encryption, decryption, and arithmetic operations intrinsic to HE.

Another intricacy of He is the inevitable introduction of “noise” in the ciphertext during its processes. As consecutive operations are conducted on the ciphertext, this noise accumulates, which might create a scenario where accurate decryption becomes

infeasible. This problem became less significant with the introduction of “bootstrapping” (Gentry, 2009). This innovation facilitated unrestricted operations on encrypted data, making schemes “bootstrappable " as they could evaluate their augmented decryption schemes.

Combination of Homomorphic Encryption and Biometric Authentication

Use of homomorphic encryption in Biometric Authentication Systems

Integrating HE into PPBA systems offers a multifold advantage. Firstly, it ensures that biometric data, being sensitive by nature, is never exposed during the authentication process. Leveraging HE allows biometric data to be processed in its encrypted form, enabling accurate authentication outcomes without revealing the underlying raw data. Such an approach addresses a major concern associated with biometric systems: the potential misuse of biometric data if compromised. Having the ability to send and operate on confidential data without revealing the plaintext, creates a scenario where organisations can use HE-equipped PPBA systems within untrusted environments, expanding their applicability.

However, this integration is not without its challenges. As stated in above sections the computational overhead associated with HE can be significantly higher than traditional methods, potentially impacting real-time applications. While Gentry’s introduction of Fully Homomorphic Encryption has diminished some of these concerns, particularly regarding noise accumulation, practical implementations still face hurdles in terms of computational efficiency. Additionally, the intricate nature of HE mandates deep mathematical comprehension, making its widespread implementation a formidable task for the average developers.

Despite this issue, the marriage of HE and PPBE holds undeniable potential. As computational resources continue to evolve and as optimisation techniques develop, the integration challenges are likely to dissolve. This paired with the rigorous research and continuous refinement, PPBA systems augmented with HE can pave the way for a new paradigm in biometric schemes.

Code Implementation

Language Suitability

Python:

- **PyCrypto and PyCryptodome:** These two python libraries are designed to use cryptographic operations. Providing functionalities for secure hashing, encryption, and decryption, among other cryptographic primitives.
- **Ease of Use:** Python is regarded as a 'beginner' coding language due to its simplicity and readability, which could be beneficial when implementing my cryptographic algorithms or integrating libraries.
- **Integration:** Python often provides systems called wrappers for other languages, allowing for integration with C/C++ libraries if needed.

C/C++:

- **OpenSSL:** Toolkits for the Transport Layer Security (TLS) and Secure Sockets Layer (SSL) protocols. OpenSSL is a general purpose library that helps give open-source implementation of both SSL and TLS protocols, it is also equipped to handle various cryptographic operations.
- **Crypto++:** A C++ library for cryptographic schemes. It is trusted by developers for its reliability and performance.
- **Performance:** C/C++ offers closer-to-hardware programming, meaning it is usually faster in its execution, especially in situations where systems are computationally intensive, like Homomorphic Encryptions.
- **Memory Control:** With C/C++ developers can have a more direct control over the memory allocation and deallocation, which could be crucial as I'll be dealing with sensitive information.

Summary:

After evaluating which language would best suit the creation of my PPBA system, I am inclined to lean towards Python, this is primarily because of its user-friendliness, clarity, and vast library support. Along with recommendations from my supervisor Python has libraries dedicated to cryptography and privacy preservation, along with its capabilities for processing biometric data. These aspects align seamlessly with the objectives of my development project.

While C/C++ has its advantages in terms of performance and memory control, its steeper learning curve and the complexities associated with its implementations could arouse potential challenges during the creation process. Taking into account my personal proficiency, coupled with the guidance from my supervisor, I believe Python is the most suitable choice for the development of my PPBA system.

Library Selection

In my journey to construct a robust Privacy-Preserving Biometric Authentication system, I faced the pivotal decision of selecting the right homomorphic encryption library. After rigorous exploration of available options, I zeroed in on TenSEAL. Tailored for Python, this library elegantly bridges the divide between homomorphic encryption techniques and practical demands of machine learning algorithms. What further solidified my confidence in TenSEAL was its foundation on Microsoft's SEAL, renowned as one of the foremost homomorphic encryption libraries. Its seamless compatibility with Python, the chosen language for my project, assures not only robustness but also a streamlined integration process.

Conclusion

Privacy-Preserving Biometric Authentication (PPBA) signifies a crucial advancement in the realm of digital security, addressing the paramount need to combine biometric accuracy with individual privacy. Integrating Homomorphic Encryption (HE) into this domain allows operations on encrypted biometric data without the necessity for decryption, ensuring the data's sanctity remains uncompromised throughout the authentication process. While HE brings groundbreaking potential, it isn't without challenges, like computational overhead and intricate mathematical complexities. Despite these challenges, as computational techniques advance, the prospect of seamless HE and PPBA integration becomes more palpable. The choice of Python for our PPBA system development, due to its adaptability and rich library support, emphasises the importance of tool selection in achieving desired outcomes. In essence, the fusion of biometrics and encryption through PPBA promises a future where digital authentication is both secure and respects individual privacy. This intersection of technology is not only timely but paramount for the ever-evolving digital landscape.

References

1. *The history of biometric authentication* (no date) *Thales Group*. Available at: https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/inspired/history-of-biometric-authentication#:~:text=The%20History%20of%20Biometric%20Authentication&text=Biometrics%20can%20be%20traced%20back,the%20work%20of%20Alphonse%20Bertillon. (Accessed: 18 October 2023).
2. University, C.G.S. *et al.* (2009) *Fully homomorphic encryption using ideal lattices: Proceedings of the forty-first annual ACM Symposium on Theory of Computing*, *ACM Conferences*. Available at: https://dl.acm.org/doi/pdf/10.1145/1536414.1536440 (Accessed: 17 October 2023).
3. *Types of biometrics* (2022) *Biometrics Institute*. Available at: https://www.biometricsinstitute.org/what-is-biometrics/types-of-biometrics/ (Accessed: 18 October 2023).
4. (No date) *Biometric authentication: A Review*. Available at: https://www.biometrie-online.net/images/stories/dossiers/generalites/International-Journal-of-u-and-e-Service-Science-and-Technology.pdf (Accessed: 18 October 2023).
5. (No date a) *Biometric authentication. types of biometric identifiers - theseus*. Available at: https://www.theseus.fi/bitstream/handle/10024/44684/Babich_Aleksandra.pdf (Accessed: 18 October 2023).
6. (No date a) *Biometrics: A tool for information security - IEEE xplore*. Available at: https://ieeexplore.ieee.org/document/1634356 (Accessed: 19 October 2023).
7. A. K. Jain and U. Uludag, "Hiding biometric data," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 11, pp. 1494-1498, Nov. 2003, doi: 10.1109/TPAMI.2003.1240122.
8. (No date a) *Biometric authentication: System security and user privacy | IEEE ...* Available at: https://ieeexplore.ieee.org/document/6353458 (Accessed: 19 October 2023).
9. Rathgeb, C. and Uhl, A. (2011) *A survey on biometric cryptosystems and cancelable biometrics - EURASIP Journal on Information Security*, *SpringerOpen*. Available at: https://jis-eurasipjournals.springeropen.com/articles/10.1186/1687-417X-2011-3 (Accessed: 19 October 2023).
10. (No date a) *[PDF] a fully homomorphic encryption scheme | semantic scholar*. Available at: https://www.semanticscholar.org/paper/A-fully-homomorphic-encryption-scheme-Gentry/5496636b7474ef68f79248de4a63dd879db55334 (Accessed: 19 October 2023).

11. Frank, M. *et al.* (2012) *Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication*, *arXiv.org*. Available at: https://arxiv.org/abs/1207.6231 (Accessed: 19 October 2023).
12. ISO/IEC (2006). Information technology — Biometric performance testing and reporting — Part 1: Principles and framework. ISO/IEC 19795-1:2006.
13. Bolle, R.M., Connell, J.H. and Ratha, N.K. (2002). The relation between the ROC curve and CMC. Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'02).
14. ElGamal, T. (1985). A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE transactions on information theory, 31(4), 469-472.
15. *The challenges of homomorphic encryption* (2022) *Atos*. Available at: https://atos.net/en/lp/cybersecurity-magazine-enter-a-new-cybersecurity-era/the-challenges-of-homomorphic-encryption (Accessed: 19 October 2023).
16. (No date a) *A secure biometric authentication scheme based on robust hashing*. Available at: https://www.semanticscholar.org/paper/A-secure-biometric-authentication-scheme-based-on-Sutcu-Sencar/b988b589ef61114a2794b28d3f0826a567c7a2bb (Accessed: 20 October 2023).
17. (No date a) *1 a survey on homomorphic encryption schemes: Theory and implementation*. Available at: https://web.eng.fiu.edu/aacar001/papers/fhe-survey.pdf (Accessed: 20 October 2023).
18. (No date a) *Bridging the gap: From biometrics to forensics - Michigan State University*. Available at: http://biometrics.cse.msu.edu/Publications/GeneralBiometrics/JainRoss_BiometricsForensics_PTRSB_15.pdf (Accessed: 20 October 2023).