Biometric Data vs. Biometric Template

1. Introduction

Password verification now accounts for more than 80 per cent of cyber breaches, according to a 2022 Verizon report [1], biometric authentication has already proven to be far more reliable than traditional authentication methods such as passwords.

The term *biometrics* is derived from the Greek word *bio*, meaning *life*; and *metric*, meaning *to measure*. The basic premise of biometric authentication is that every person can be accurately identified by intrinsic physical or behavioral traits. In biometric recognition system, once the biometric traits of a person (such as handprints, facial features) are captured by the device, a formula or algorithm is used to compute a mathematical representation from the data which is then matched against the person's reference. Such mathematical format of biometric data is called a *biometric template*. As illustrated in figure 1, this representation or template is a high dimension vector of real numbers computed from an algorithm and a small difference between incoming and reference indicates a match.

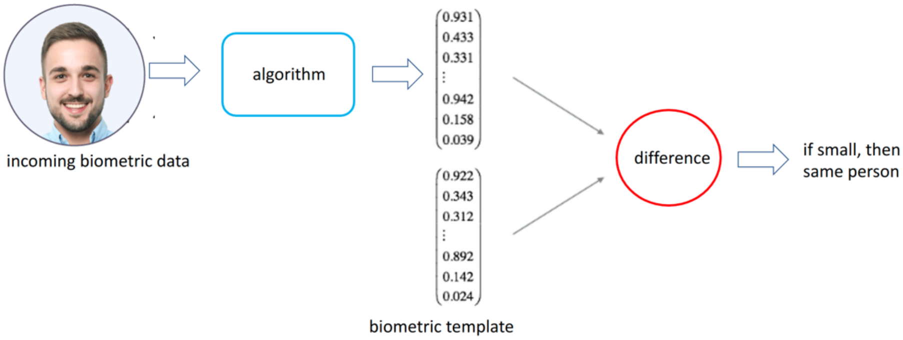

Figure 1. A biometric template is a set of numbers for matching to a person.

No photos or media files are needed for future comparisons, so they are **automatically and permanently deleted** after the biometric template is generated at enrollment – specifically in ParallelWallet, the deletion includes images of the user's face, palm, and voice samples – *the source code for this specific process will be publicly posted on our GitHub upon the product launch*. Therefore, the general public's fear that, in the event of biometric data being compromised, the hacker would be able to reuse the person's original face, handprint or whatever the biometric happens to be, is unfounded. In order to reuse the templates generated by ParallelWallet, a hacker must first have a compatible deep-learning model, then crack the encryption and re-assemble the files properly, and finally de-scramble the proper ordering (described below), which is a non-trivial and daunting process.

In the case of the handprint template for the palm recognition in ParallelWallet, we also deliberately exclude all fingerprints for privacy protection. Only the area from the palm up to the middle knuckles is used for generating the template.

2. Anti-spoofing and multi-biometric templates

The matching of biometric template is only one component in the identity authentication process. It addresses the first question of "Is this person who they claim to be?"; followed by another component called "*anti-spoofing*" to address the further question of "Is this a real person?". Both checks must pass before the user is admitted (see figure 2). Therefore, a hacker's attempt to pass authentication by using a photo of your face or handprints will fail.

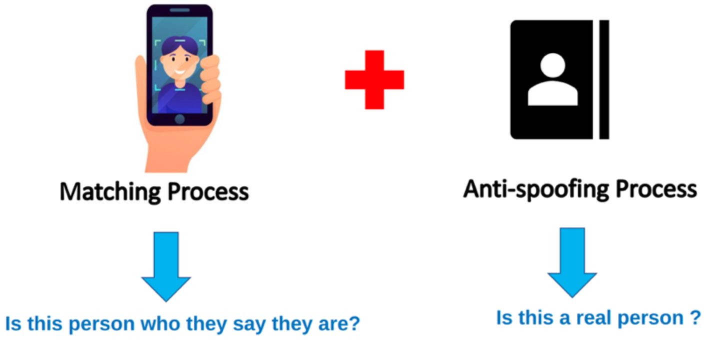

Figure 2. Two processes are checked: "Is this the same person?" "Is this a real person?"

Users who want maximum security can fully utilize the multi-biometric authentication in ParallelWallet, it supports four aspects of identity authentication in a single application, which is unique and unseen in other biometric identity products or services in the market today: "who you are (biometrics)", "is this a live person (anti-spoofing)", "what you know (secret passphrase in voiceprint)", and "what you possess (mobile device in use)".

3. Processing and storing biometric templates safely

The security of your biometric data in ParallelWallet's authentication system is not limited to this level only. First, each web request to the biometric server is issued a time-limited token per user session which expires after a certain period (e.g., 10 minutes) to reduce data security risk in action ("*data in action*"). Second, all data is encrypted during internal transmission and storage ("*data in transit*" and "*data at rest*"). Third, the biometric template is stored in a format that meets the requirements Biometric data and privacy laws [2] such as the GDPR[1], to protect user's privacy, even in the unlikely event of template theft.

To achieve this third safety measure, each biometric template (which consists of a vector of real numbers) is further scrambled by a secret internal algorithm and split into multiple pieces before being encrypted and stored. We make sure the scrambling method destroys all correlation to the user's biometric data and, subsequently, devoid of any private information. The splitting into multiple files allows pieces of a template to be placed on separate servers or use different encryption keys thus further reducing the risk of full extraction. In the unlikely event of a data breach, we can simply revoke the old scrambling process, abandon the existing templates, and regenerate new ones using a new version of scrambling.

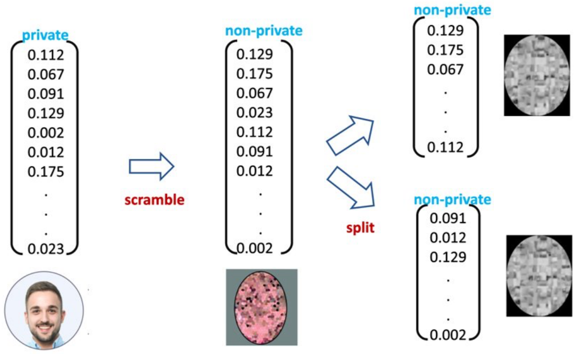

Figure 3. Each biometric template is scrambled by an internal algorithm and then split into numerous pieces to preserve privacy. A 128-dimension template vector will generate 128! combinations for a brute force attack which is approximately 3.8562048236258E+215 times. Vector dimensions of 256 or 512 are not uncommon.

In summary, the ParallelWallet biometric authentication allows users to use biometrics safely and privately leveraging upon the various protective measures described above.

References

[1] Verizon, 2020. Data Breach Investigations Report. Available at:
https://www.verizon.com/business/resources/reports/2022/dbir/2022-data-breach-investigations-report-dbir.pdf

[2] Thales, 2021. Biometric data and privacy laws (GDPR, CCPA/CPRA). Available at:
https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/biometrics/biometric-data

[3] Norwegian University of Science and Technology, 2021. Biometric Privacy and Information Protection. Available at:
https://www.christoph-busch.de/files/Busch-Privacy-210510.pdf

[1] The GDPR is the first and most widely accepted legal ordinance that specifically addresses the handling of biometrics data for privacy protection. The stored form of biometric template is expected to be (i) accurate, (ii) un-invertible (so original data cannot be reverse engineered from the template), (iii) revocable/cancelable (in case of a breach) and (iv) unlinkable from application to application [3].