# Managing Private Credentials by Privacy-Preserving Biometrics

Bian Yang and Guoqiang Li

Norwegian University of Science and Technology (NTNU), 2815 Gjøvik, Norway
{bian.yang;guoqiang.li}@ntnu.no

**Abstract.** We investigate in this paper the need to managing a user's private credentials using privacy-preserving biometrics, define several private credential management work models under different trust models between a user and an external party. A general pipeline using privacy-preserving biometrics for private credential management is proposed to achieve the purpose of biometric template protection, biometric-secret binding, and biometric recognition accuracy performance improvement. The proposed scheme was implemented and tested in the European CIP project PIDaaS, and demonstrated advantages in privacy preservation and accuracy performance preservation.

**Keywords:** private identity, privacy-preserving biometrics, biometric template protection, biometric-secret binding, user-centric identity management.

## 1 Introduction

An identity authentication mechanism enables a service provider to distinguish their customers and customize the service for each of them. However, in most services, managing users identities increases the business operational burden and also the risk of data leakage. As the GDPR strengthens the data subjects' rights to their personal data and specifies the penalty on data breach, service providers can be motivated to outsource customer identity management to a professional party in order to reduce the risk of data breach. Identity management outsourcing models can include identity federation and Single-Sign-On (SSO) (*e.g.*, the Google Identity Platform [1] and the Facebook Login [2]), a cross-service identity platform (*e.g.*, OpenID [3] and bankID [4]), and claimed-based identity management schemes [5] [6].

While service providers have many options (as mentioned above) to outsource the identity management task, users usually have to manage (memorize, take a note, save a file, *etc*.) by themselves their identities for authentication (*e.g.*, account name, identification number, password, PIN, private key, *etc*.) – called "private credential" in this paper. As there are increasing private credentials for an ordinary user to manage, the user has a growing need to outsource credentials' management to a professional party. This party can be either a software such as a password manager, or an organization delegated to compute in an authentication protocol. A password manager with a master user account and secret is for instance a typical way to manage credentials.

When it comes to the possibility using biometrics to replace a master password to manage a credential manager [7], the configuration for outsourcing to a private credential manager (PCM) can be complicated. Using biometrics implies processing the biometric sample, extracting and protecting the derived feature, and securely storing the generated biometric template. All these burden a user in terms of computation and security. In addition, the user may still need to manage a master secret (*e.g.*, a master password to the PCM, a secret key for generating a protected template, *etc.*), or other supplementary data (SD) that could be needed in operating a PCM. The ISO/IEC JTC1 24745 on Biometric Information Protection [8] presented a general model for processing a plain biometric feature ***b*** and generating a protected template (PT), which includes a pseudonymous identifier (PI) for direct comparison and an auxiliary data (AD) for reconstructing a new PI for comparison. What data out of {***b***, SD, PI, AD} and the associated computations can be delegated to a PCM may depend on various considerations (*e.g.*, efficiency, reliability, cost, security, trust model, law compliance, *etc.*) among which the trust model can be decisive.

One noticeable step towards the concept of biometrics-enabled private credential management was made by the FIDO alliance [9], whose UAF standard provides a general way to binding and unlocking a service-specific private key for authenticating the user to a service provider via the FIDO Authenticator placed in the device. The drawbacks of the FIDO solution include (1) it is a device-centric solution, *i.e.*, the service-specific private key is generated per service provider, per device, and per user account, and the biometric verification takes place at the device. Upon a device loss a user has to revoke the certificate and the private key associated with the lost device, and create a new registration including enrolling her/his biometrics on the new device. This implies hardly any portability, which is not compatible with the concept of the claimed-based identity management. (2) it is not specified in FIDO UAF how a biometric template is stored and how a service-specific private key can be unlocked from a biometric verification. The security in storing the biometric template and the private key depends highly on the device's hardware and software environment for protecting these data. The variance in devices and FIDO UAF authenticator's implementation increases the complexity of configuring data protection on a specific device.

Instead of following the device-centric concept, we propose in this paper a user-centric approach to managing private credentials by privacy-preserving biometrics. In Section 2, we propose three typical work modes in which a biometric private credential manager (PCM) can be configured to manage credentials. In Section 3, we propose a privacy-preserving biometric-secret binding scheme. Section 4 gives the performance testing results of the proposed scheme, and Section 5 concludes this paper.

## 2 Configuring a Biometric Private Credential Manager

We denote in this paper the service provider as SP, a user who wants to authenticate her/him with a SP as User, and the identity provider as IdP. The identity data held and managed by a User for authentication is a Credential, and the data held and managed by an IdP or a SP to attest and ascertain the User's identity claim is a Registry. A SP can outsource the identity authentication computation or even decision making to an IdP as a party trusted by the SP, which is always assumed in this paper. Similarly, a User may reply solely on a software (or with hardware additionally) denoted as a Private Credential Manager Client (PCMC) for managing their credentials, or outsource data storage and part of computation required for authentication to an independent party denoted as a Private Credential Manager Server (PCMS). From a User's perspective, how to configure the data {biometric feature $\boldsymbol{b}$, supplementary data $SD$, pseudonymous identifier $PI$, auxiliary data $AD$} (explained in Section 1) and the associated computations among different parties can be varied.

One decisive factor is how much trust a User has on PCMS so that (s)he can decide whether to outsource to PCMS the storage of $AD$ (*e.g.*, a protected biometric template), or instead, the storage, part of template generation computation, and even the cryptographic authentication process interacted with IdP. We define the following three work modes of a PCM based on a User' different levels of trust on a PCMS. Note that the defined work modes do not represent all possible configurations among the parties. Instead, they are assumed the typical ones from the User perspective.

Work Mode I (local computation and storage): as showed in Fig. 1(a), the User relies solely on the PCMC to take as input $\boldsymbol{b}$, $SD_1$ (*e.g.*, a master password or PIN), and $SD_2$ (*e.g.*, a salt managed by IdP), and generate and store $AD$ locally on the User's device. Together with $AD$ is generated $PI$ which can be used as the secret for a cryptographic authentication protocol between PCMC and IdP. This work mode can be deemed as a biometric version of a local password manager or a FIDO UAF authenticator. The User has to fully trust her/his device and software on it.

Work Mode II (local computation and outsourced storage): as showed in Fig. 1(b), the User relies on the PCMC to take as input $\boldsymbol{b}$, $SD_1$, and $SD_2$, and generate $AD$. The $AD$ is stored in PCMS and will be retrieved by PCMC for generating a new $PI$ used as the secret for a cryptographic authentication between PCMC and IdP. This work mode outsources the storage of $AD$ to PCMS, and therefore the User has to trust the PCMS on its capacity of properly protecting an $AD$ from leakage if this $AD$ should be kept confidential. Otherwise the User should use encryption or a biometric template protection scheme [10] to ensure $AD$ does not reveal any information about $\boldsymbol{b}$.

Work Mode III (distributed computation and outsourced storage): as showed in Fig. 1(c), the User relies on the PCMC to take as input $\boldsymbol{b}$, $SD_1$, and $SD_2$. Instead of generating $AD$ directly, the PCMC can perform a lightweight protection $f(\boldsymbol{b}, SD_1, SD_2)$ on $\boldsymbol{b}$ and send the partially protected $\boldsymbol{b}$ to PCMS. Then the PCMS will continue the biometric feature protection, verification, and the cryptographic authentication with IdP. Splitting up the work effort between PCMC and PCMS can ensure that (1) PCMC's computational workload be shared by PCMS, when PCMC is operated on a power-constrained mobile device; (2) PCMS provide better protection for $AD$ and $PI$ when for any reasons the PCMS is more trusted than the PCMC; (3) PCMC provide an extra layer of protection for $\boldsymbol{b}$ when the PCMS is not fully trusted.

Table 1 summarizes the three work modes qualitatively evaluated in different criteria. The proposed privacy-preserving scheme in Section 3 can make the most of the Work Mode III when both storage and part of computation are outsourced to a less trusted party as PCMS.

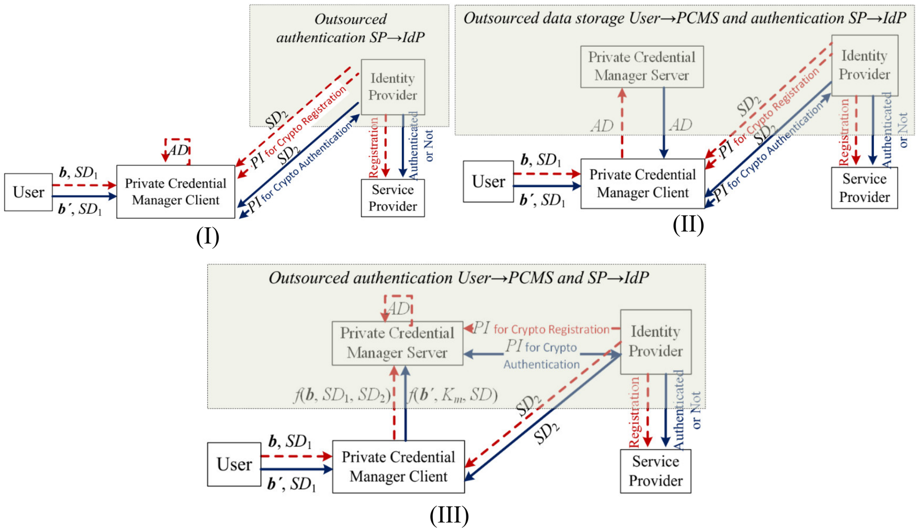

Outsourced authentication SP→IdP
Identity Provider
AD
b, SD1
b′, SD1
User
Private Credential Manager Client
SD2
PI for Crypto Registration
SD2
PI for Crypto Authentication
Registration
Authenticated or Not
Service Provider
(I)
Outsourced data storage User→PCMS and authentication SP→IdP
Private Credential Manager Server
AD
AD
Identity Provider
User
b, SD1
b′, SD1
Private Credential Manager Client
SD2
PI for Crypto Registration
SD2
PI for Crypto Authentication
Registration
Authenticated or Not
Service Provider
(II)
Outsourced authentication User→PCMS and SP→IdP
AD
Private Credential Manager Server
PI for Crypto Registration
PI for Crypto Authentication
Identity Provider
f(b, SD1, SD2)
f(b′, Km, SD)
SD2
SD2
Registration
Authenticated or Not
User
b, SD1
b′, SD1
Private Credential Manager Client
Service Provider
(III)

**Fig. 1.** Work modes for a biometric private credential manager

**Table 1.** Comparison among work modes of a Private Credential Manager.

| Performance of a PCM from User's perspective | Mode I | Mode II | Mode III |
|---|---|---|---|
| Computational complexity in local device | High | Medium | Low |
| Cost in outsourcing | N/A | Low | High |
| Credential portability | Low | High | High |
| Trust required on local device and PCMC | High | Medium to High | Medium |
| Trust required on PCMS | N/A | Low | High |

# 3 The Construction of a Biometric-Secret Binding Scheme

## 3.1 Conventional Biometric-Secret Binding Schemes

Biometric-secret binding schemes can combine a biometric feature and a cryptographic secret during enrolment and release the secret for authentication during verification when the probed biometric feature is close enough to the one used in enrolment. A general framework of biometric-secret binding was presented in the Fig.1 in [11]. Typical construction methods for such a scheme include fuzzy commitment [12], fuzzy vault [13], and secure sketch [14] (note that QIM was used in [14][11] to encode secret bits by choosing quantizers, which should be distinguished from the other type of secure sketch [15][16] which generates secret bits from biometric features). Due to lack of effective feature processing steps, such biometric-secret binding schemes are known for their distinct biometric recognition accuracy degradation.

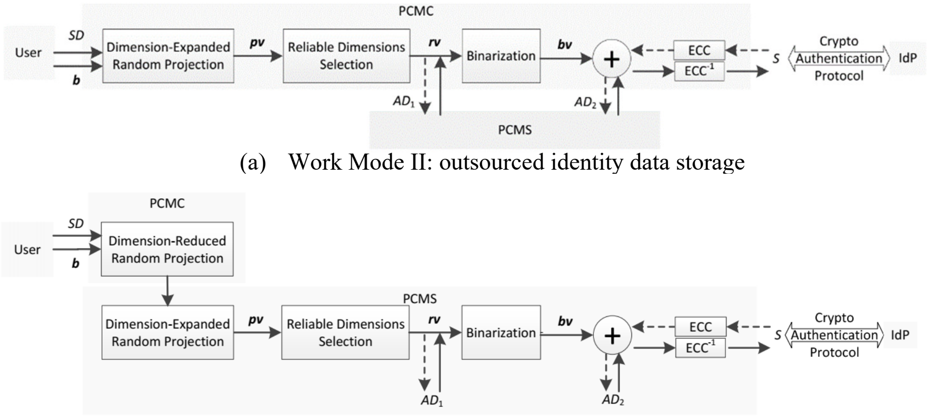
Unlike conventional fuzzy commitment schemes with compromised biometric recognition accuracy and the security concern of key-inversion (*i.e.*, deriving plain biometric feature from the leaked secret), we propose in Section 3.2 a new feature transformation with dimension-expanded/reduced random projection and reliable dimensions selection in order to (1) preserve the biometric recognition accuracy; (2) enhance the protection of the biometric information beyond that of the secret so that an key-inversion attack will hardly work; and (3) shield the plain biometric feature securely from a delegated authentication operator (*e.g.*, the PCMS) in Work Mode III.

### 3.2 The Proposed Privacy-Preserving Biometric-Secret Binding Scheme

The proposed privacy-preserving biometric-secret binding scheme consists of four steps in sequence: random projection, reliable dimensions selection, binarization, and fuzzy commitment. Suppose a User possesses her/his biometric feature vector ***b***, the supplementary data *SD* (including both $SD_1$ under her/his control and $SD_2$ which can be retrieved from IdP upon a request from the User), and a secret *S* for binding.

Enrolment:

**Step 1: Dimension-Reduced and/or Expanded Random Projection.** Each ***b*** can be transformed to a vector ***pv*** via random projection which largely preserves the distance. A random projection can be a dimension-reduced one (*i.e.*, a surjection) so that ***b*** is hard to precisely recover from ***pv*** due to information loss. This distance-preserving property makes random projection an attractive solution [17] to biometric template protection. Random projection can be also used to expand the dimension of ***b*** in order to obtain better biometric comparison accuracy [18][19] when using simple quantization (*e.g.*, thresholding by sign) to binarize ***pv***. The dimension-reduced and the dimension-expanded random projections can be also used in tandem in order to achieve both privacy protection (irreversibility defined in [20]) for ***b*** and maximized recognition accuracy from ***pv***. In practice, it can be at the User's decision either or both of the two can be adopted in the proposed pipeline, depending on the need of the trust on the outsourced PCMS or the recognition accuracy performance. **Step 2: Reliable Dimensions Selection.** The expanded random projection provides a rich set of projected dimensions among which the most reliable dimensions can be selected to construct a vector ***rv*** as a reliable representation of ***b***. The reliability can be defined in the sense that a dimension $rv_i$ ($1 \leq i \leq N$, where *N* is the dimension of ***pv***) has an as-small-as-possible intra-class distance (from an expectation value of $rv_i$ of samples from the same biometric characteristic) while an as-large-as-possible inter-class distance (from the same dimension of other users). In practice, multiple samples are needed from the same biometric characteristic to estimate the expectation value in each dimension, which can be done during enrolment. A separate dataset is needed to form an "imposter set" to calculate the inter-class distances between an enrolled ***rv*** and those projected vectors generated from the imposter set. To rank all the dimensions $rv_i$ ($1 \leq i \leq N$) in reliability, simple metrics such as Equal Error Rate can be adopted to measure the dis-similarity of the two distance sets' distribution. The indices of the selected dimensions are saved as $AD_1$ locally (Work Mode I) or in PCMS (Work Mode II and III). **Step 3: Binarization.** A binarization step follows the reliable dimensions selection in order to convert the selected reliable dimensions (*i.e.* vector ***rv***) to a binary representation, denoted as a vector ***bv***, to be used for the next step of


PCMC
User
SD
b
Dimension-Expanded Random Projection
pv
Reliable Dimensions Selection
rv
Binarization
bv
+
ECC
ECC-1
S
Crypto Authentication Protocol
IdP
AD1
AD2
PCMS
PCMC
User
SD
b
Dimension-Reduced Random Projection
PCMS
Dimension-Expanded Random Projection
pv
Reliable Dimensions Selection
rv
Binarization
bv
+
ECC
ECC-1
S
Crypto Authentication Protocol
IdP
AD1
AD2

(a) Work Mode II: outsourced identity data storage

(b) Work Mode III: outsourced authentication

**Fig.2** Proposed privacy-preserving biometric-secret binding pipeline in Work Mode II and III

secret binding. A binarization method (*e.g.*, simple thresholding by mean value or sign, or a unary representation) can generate a vector that is suitable for Hamming distance calculation, implying that all generated binary bits should be approximately equally weighted and distributed. **Step 4: Secret Binding - Fuzzy Commitment.** Assuming a secret *S* (*e.g.*, password or a private key) has been created for a cryptographic authentication protocol between the User (PCMC/PCMS) and the IdP, we can use a secret binding scheme to combine *S* and ***bv*** into an $AD_2$ stored locally (Work Mode I) or in PCMS (Work Mode II and III). We use the fuzzy commitment scheme [12] to achieve this purpose. This enrolment process takes as input {***b***, $SD_1$, $SD_2$, *S*} and outputs the protected template {$AD_1$, $AD_2$}, whose breach would have only limited risk in privacy breach regarding ***b***, if not all three data {$SD_1$, $SD_2$, *S*} are leaked.

Verification:

**Step 1: Dimension-Reduced and/or Expanded Random Projection.** During verification, a biometric feature vector ***b'*** together with {$SD_1$, $SD_2$} is input to the pipeline, and the step is exactly same as the Step 1 during enrolment. **Step 2: Reliable Dimensions Selection.** The saved $AD_1$ recording selected indices is retrieved and supplied to the projected vector ***pv'*** derived from ***b'*** to form a ***rv'*** as a reliable representation of ***b'***. **Step 3: Binarization.** The same method as in enrolment is used to obtain a binary representation vector ***bv'*** from ***rv'***. **Step 4: Secret Releasing - Fuzzy Commitment.** Via the secret releasing step in fuzzy commitment, an error-correction decoded result *S'* is released for the subsequent cryptographic authentication protocol.

Depending on the trust a User has on an independent PCMS, the User can choose to complete the Step 1-4 in the three different work modes. In the Work Mode I, all the 4 steps are completed in the local device (PCMC) that is fit for the scenario where the User has low trust level on an external party besides IdP. In the Work Mode II and III, the User has medium and high trust on a PCMS respectively, and therefore storage and even part of computation in protected template generation can be outsourced to a PCMS. Fig.2 (a) and (b) show the pipeline described above in the Work Mode II and III respectively with different steps and data managed by different parties. The underlying rationale for the difference is that a User may add an extra protection, in addition to fuzzy commitment, for ***b*** via a dimension-reduced random projection step prior to sending ***b*** to PCMS. The stored protected biometric template $\{AD_1, AD_2\}$ can be safely stored locally or outsourced with little concern in privacy leakage of ***b***. Via changing the random matrix for projecting ***b*** and the secret $S$ used by fuzzy commitment, unlinkable $\{AD_1, AD_2\}$ can be derived from the same biometric characteristic.

## 4 Performance Testing

We tested the proposed scheme in the pilots developed under the EU-CIP project PIDaaS (www.pidaas.eu). Two biometric modalities – face and voice – with the COTS SDK from Viulib [21] and Alize [22] respectively, were adopted in the pilots. Both SDKs generate fixed-length plain biometric feature vector ***b***. In total 47 participants contributed to the test datasets. 3 face samples were used for enrolment and generating $AD_1$, and another 34 face samples were used as probes. 5 voice samples (in each sample a random set of 5 digits were spoken) with / without environmental noises were used for enrolment and generating $AD_1$, and another 6 voice samples were used as probes. If a probe vector is sufficiently close to the enrolled vector, the secret $S$ could be successfully released. As the authentication conclusion is binary, we got only one performance point in terms of False Match Rate (FMR) and False Non-Match Rate (FNMR) instead of a continuous Detection Error Tradeoff curve or an Equal Error Rate (EER). The performance reported here from voice recognition were generated from the "imposter scenario" where all feature vectors included in inter-class distance calculation were derived from the same set of 5 digits spoken. Table 2 presents the performance from two representative sets of 5 digits, and Table 3 presents the performance from the face case. Though the accuracy performances were calculated from small-scale datasets, we observed that the biometric probes can be matched with well-preserved accuracy universally. The non-degradation in performance could be attributed to the dimension-expanded random projection and the reliable dimension selection steps that made the most of biometric features.

**Table 2.** Comparison in recognition accuracy – voice with 2 different sets of 5 digits spoken

| Voice recognition method | FMR(%) | FNMR(%) | EER(%) |
|---|---|---|---|
| Plain voice without protection | 0/0 | 2.4/9.5 | 0.2/1.2 |
| The proposed privacy-preserving voice | 0/0 | 3.5/5.9 | n/a/ |

**Table 3.** Comparison in recognition accuracy – face

| Face recognition method | FMR(%) | FNMR(%) | EER(%) |
|---|---|---|---|
| Plain face without protection | 0 | 33.3 | 1.0 |
| The proposed privacy-preserving face | 0 | 9.3 | n/a |

## 5 Conclusion

Three typical work modes and a general pipeline for constructing the privacy-preserving biometrics based credential management were in this paper. The proposed biometric-secret binding scheme has the following advantages: (1) biometric modality and feature agnostic; (2) recognition accuracy performance preserving; (3) privacy-preserving (irreversibility and unlinkability) biometric templates; (4) extra protection for the biometric information when the authentication function is outsourced.

## References

1. Google Identity Platform: https://developers.google.com/identity/.
2. Facebook Login: https://developers.facebook.com/docs/facebook-login/.
3. OpenID: https://zh.wikipedia.org/wiki/OpenID.
4. bankID: https://www.bankid.no/en/company/.
5. Claim-Based Identity: https://en.wikipedia.org/wiki/Claims-based_identity.
6. Alrodhan, W., Mitchell, C.: Enhancing User Authentication in Claim-Based Identity Management. In: 2010 Int. Sym. on Collaborative Technologies and Systems, pp.75-83. (2010)
7. Yang, B., Chu, H., Li, G., Petrovic, S., Busch, C.: Cloud Password Manager Using Privacy-Preserved Biometrics. In: Proc. of 2014 IEEE Int. Conf. on Cloud Engineering, 2014.
8. ISO/IEC 24745: Biometric information protection, 2011.
9. FIDO Alliance: https://fidoalliance.org/.
10. Nandakumar, K., Jain, A.: Biometric Template Protection: Bridging the Performance Gap Between Theory and Practice. IEEE Signal Processing Magazine, 32(5), pp.88-100, 2015.
11. Bui, F., Martin, K., Lu, H., K. Plataniotis, and Hatzinakos, D.: Fuzzy Key Binding Strategies Based on Quantization Index Modulation (QIM) for Biometric Encryption (BE) Applications. IEEE Trans. on Information Forensics and Security, 5(1), pp.118-132, 2010.
12. Juels, A., Wattenberg, M.: A Fuzzy Commitment Scheme. In: Proc. of the 6th ACM Conference on Computer and Communications Security, pp.28-36, 1999.
13. Juels, A., Sudan, M.: A Fuzzy Vault Scheme. Designs, Codes and Crypto., 38(2), 2006.
14. Buhan, I., Doumen, J., Hartel, P., and Veldhuis, R.: Constructing Practical Fuzzy Extractors Using QIM Centre for Telematics and Information Technology, University of Twente, Enschede, Tech. Rep. TR-CTIT-07-52, pp.1381-3625, 2007.
15. Sutcu, Y., Li, Q., and Memon, N.:Protecting Biometric Templates with Sketch: Theory and Practice. IEEE Trans. Inf. Forensics Security, 2(3), pp. 503–512, 2007.
16. Dodis, Y., Ostrovsky, R., Reyzin, L., and Smith, A.: Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. SIAM J. Comp., 8(1), 2008.
17. Teoh, A., Ngo, D., and Goh A. Biohashing: Two Factor Authentication Featuring Fingerprint Data and Tokenised Random Number. Patt. Recog., 37(11), pp.2245–2255, 2004.
18. Yang, B., Busch, C., Gafurov, D., Bours, P.: Renewable Minutiae Templates with Tunable Size and Security. ICPR, pp.878-881, 2010.
19. Yang, B., H., Daniel, Simoens, K., Busch, C.: Dynamic Random Projection for Biometric Template Protection. IEEE BTAS, pp.1-7, 2010.
20. Simoens, K., Yang, B., Zhou, X., Beato, F., Busch C., Newton, E., Preneel, B.: Criteria towards Metrics for Benchmarking Template Protection Algorithms. ICB, 2012.
21. http://www.viulib.org
22. https://github.com/ALIZE-Speaker-Recognition