SECURE BIOMETRICS

Marten van Dijk, Pim Tuyls

Philips Research Laboratories
Prof. Holstlaan 4, AA 5656 Eindhoven, The Netherlands
email: marten@mit.edu, pim.tuyls@philips.com

ABSTRACT

In this paper, we extend the information theoretic secure constructions for biometrics to the computational setting. Based on semantically secure encryption, we introduce robust, fully private and secure biometric key distillation and verification. Our model incorporates an adversary with side information who has access to a database with reference information. Even though our schemes are based on a master key, no master key needs to be stored in biometric sensors. In our scheme it is possible to derive a polynomial number of keys from a single biometric and we show how to renew keys in a secure and private way without additional interaction with the user. Previous work considers unconditional secure key distillation which can at most reach partial (information theoretic) privacy and which can only lead to a small number of keys for each biometric.

1. INTRODUCTION

Biometrics identify/authenticate people on what they are rather than on what they have (tokens) or what they know (passwords). Since biometric properties can not be lost or forgotten in contrast to tokens and passwords, they offer an attractive and convenient alternative to identify and authenticate people. During the *enrollment* phase the biometric of a person is measured and a derived template is encoded in reference information and stored in a database. During the *verification* phase, a verification device measures a biometric, retrieves the corresponding reference information from the database and performs a fuzzy match. The database is publicly accessible by each verification device.

The main risks involved in using biometrics are: i) Biometrics contain sensitive information about people [B, P65]. ii) Once compromised, they are compromised forever and can not be reissued [S99]. iii) They can be used to perform cross-matching between databases and to track peoples behaviour. iv) Many biometric identifiers can be forged based on template information [PK00, B02, MMJ$^+$03]. So, unprotected biometric reference information in the database leads to privacy risks; i.e. they leak information on the biometric. This problem received recently a lot of attention [JS02, TG04, LT03, DRS04, JW99, SRS$^+$98].

In current template protection schemes [DRS04, LT03, TG04] the reference information in the database consists of *helper data*, which is used to correct measurement noise and derive a key, and a cryptographic hash of the key. The cryptographic hash is used to verify the derived key. The helper data is designed such that it reveals no information about the distilled key and a minimal amount of information on the biometric template. The helper data is stored in the database and inevitably leaks some (Shannon) information about the biometric template. This information can be used by an adversary to guess the biometric template with an improved probability. For example, if the Hamming distance is used, then the helper data contains linear combinations of the bits representing the biometric template. Hence, by guessing a solution of this system of linear equations (in polynomial time), an adversary improves his probability of successful impersonation. Therefore, current schemes achieve only partial privacy of biometric templates.

In this paper we show how to construct reference information which protects and maintains the full privacy of biometric templates, even in the presence of an adversary with side information (e.g. a fingerprint captured from a glass) while still leaking no information on the distilled keys. Given the side information, we are interested in the adversaries probability of successfully guessing information about the biometric and its distilled keys. Full privacy means that an adversary is not able to improve this probability by additionally using the reference information; i.e. the publicly available reference information from the database is not useful for him.

In order to achieve full privacy we rely on semantically secure encryption. Even though this requires a secret master key, the verification devices in our schemes do not need to store any master key. In our proposed computational secure setting it is possible to derive a polynomial number of keys securely and privately from a single biometric. We show additionally how to renew keys in a secure and private way without additional interaction with the user. We prove that the current schemes [DRS04, TG04] only allow a limited number of keys per biometric.

2. MODEL AND CONTRIBUTIONS

The basic model consists of users (whose biometric is measured), multiple verification devices (VDs), a single certification authority (CA), and an adversary.

Enrollment: For each user Alice, the CA measures a biometric and encodes its template (or a feature extracted from the template) together with a randomly distilled key into reference information. The CA stores the reference information in a publicly accessible database. We assume that the CA is

trustworthy and follows the protocols as required; in particular it is assumed that she will not leak any information besides what is needed in the database. This model captures key distillation from biometrics [UPP+04, DRS04, SRS+98, LT03] which is of interest as it allows to bind content to identities instead of to devices. Our model also captures the encoding of a random or uniformly distributed key rather than a key derived from the biometric.

By $J(x;k)$ we denote the *reference information* corresponding to Alice's enrolled biometric template[1] $x \leftarrow X$ and the key $k \leftarrow K$ associated with Alice. Throughout the paper, we denote random variables by capital letters and the corresponding values by small sized letters.

Verification: Alice wishes to identify herself to a VD such that the VD can reconstruct the key k from the reference information $J(x;k)$ stored in Alice's database entry. Alice presents her biometric to the VD. Because of measurement noise, the VD measures a biometric template $y \leftarrow Y$ which differs from the originally enrolled template $x \leftarrow X$ (the measurement is modeled as a noisy channel [GT04]). After obtaining the reference information $J(x;k)$ at Alice's database entry, the VD executes a $poly(|X|)$ time function $G(y, J(x;k))$ which outputs k if $d(y,x) \leq \delta$ and outputs a ? otherwise. The function G is called a *key-extractor*.

Throughout the paper, VDs measure and verify biometrics. In current literature, these functionalities are split into a sensor, which measures the biometric, and a device, which verifies the biometric. This requires a secure and authentic channel between the sensor and device to protect the privacy of measured biometric templates. In this sense, the sensor, device, together with the secure and authentic channel form a secure tamper evident unit. This is what is called a VD in this paper. We assume that VDs are trusted and do not leak measured biometric templates or distilled keys to the outside world, that is, their software does not contain security flaws. We also assume that the verification devices are cheap and can therefore not be tamper-resistant. This means an adversary can tamper with a VD to read out its key material (after which the VD cannot be used to measure biometrics since the VD is tamper-evident). In particular, this may reveal a master key common to all VDs. Hence, *biometric verification schemes have to be designed such that the VDs do not need/store a master-key.*

The communication between the verification device and CA's database, is over a public (untrusted) network and is vulnerable to man-in-the-middle attacks. This means that *biometric verification schemes have to be designed such that the authenticity/origin of publicly communicated messages is verified.* This does not prevent database entries to be replaced with older ones which requires mechanisms based on memory integrity checking [MvOV96, p. 466-468]. For example, the database entries form a hash tree whose root is signed together with a timestamp, which is updated on a regular basis (e.g. every day). To retrieve a database entry, the entries of the whole path from the database entry to the root together with their direct siblings is read out. The VD checks the hashes, verifies the signature of the root, and checks whether its timestamp is up to date. This detects whether the retrieved entries in the database are the most current ones. In this paper we assume that the database is protected by a memory integrity checking mechanism.

[1] X may also denote the result after feature extraction from the enrolled biometric template.

Adversary: If the transmitted messages are authenticated and if the VDs do not store a common master key, then an adversary may only hope to be able to fake Alice's biometric template $x \leftarrow X$ based on side information and the reference information stored in the database. By $z \leftarrow Z$ we model the adversaries side information as a noisy version of $x \leftarrow X$. The random variables X, Y, and Z are correlated, their joint distribution is denoted by $\mathbb{P}_{X,Y,Z}$.

We define robustness and reliability of (G, J) pairs to deal with two important performance parameters of biometric systems: the False Rejection Rate (FRR) and the False Acceptance Rate (FAR). The FRR gives the probability that an honest user Alice is refused and the FAR gives the probability that impersonation by an adversary succeeds.

We may reformulate the basic model as follows. The VD receives y over a noisy channel $X \rightarrow Y$ and the adversary receives z over a noisy channel $X \rightarrow Z$. Both channels are correlated and characterized by the joint conditional distribution $\mathbb{P}_{Y,Z|X}$. By means of the public database, the CA transmits a public message (the reference information) to VD, which is also received by the adversary. By using the public message, the VD corrects the noise in y and distills a secret key. Maximized over $\mathbb{P}_X$, the maximal rate at which the CA and VD are able to agree on a secret key is equal to the forward key-capacity and is equal to the secrecy capacity without public communication $C_s(\mathbb{P}_{Y,Z|X})$ [AC93, M93].

Contributions: We analyze the security and privacy provided by (G, J) pairs. For current schemes in the information theoretic setting:

1. We show that given $\mathbb{P}_{X,Y}$ and a parameter m a fuzzy extractor can be used to construct a robust pair (G, J) which is secure against adversaries with side information characterized by $\mathbb{P}_{Z|X,Y}$ with $E_{z \leftarrow Z}(\mathsf{H}_\infty(X|Z = z))$ at least approximately m. The constructed pair is not secure for all $\mathbb{P}_{Z|X,Y}$.
2. We prove that for a given biometric template $x \leftarrow X$, at most $\approx \mathsf{H}(X|Z)/\mathsf{H}(K)$ different keys $k_i \leftarrow K$ can be encoded in reference information $J(x;k_i)$ without compromising the security of the k_i's. This means that the key corresponding to a given biometric can only be renewed at most $\approx \mathsf{H}(X|Z)/\mathsf{H}(K)$ times.
3. As discussed in the introduction, only partial privacy can be achieved [LT03].

In order to overcome these draw backs we propose computational secure and private solutions. This requires the CA to manage a master key s. We assume that there exists a ppt (master) key generating algorithm M which on input 1^n, where n is the security pa-

rameter, outputs a master public-secret key pair (p, s); $(p, s) \leftarrow M(1^n)$. Both the function G and the algorithm J depend on (p, s). We indicate this by using the subscripts p and s in J_p and G_s.

1. We show that given $\mathbb{P}_{X,Y}$ semantical secure public key encryption can be used to construct a robust triple (M, G, J) which is secure and private for all adversarial side information $\mathbb{P}_{Z|X,Y}$.
2. We prove that for each biometric the corresponding key can be $poly(n)$ times renewed without compromising the security of the keys or the privacy of the biometric.
3. Full privacy is obtained.

In the basic model the VD uses G_s, hence, it needs to store the master key s which we want to avoid. For this reason we extend our model by trusted parties (TPs). We assume that the TPs are trusted and store and do not leak the master key s. This allows the VD to outsource part of its computations, which is necessary since it is not allowed to store the master-key s. To minimize the trusted computing base we implement each TP by multiple secure servers, each knowing only a share of the master key s. We introduce homomorphic encoder-decoders and combine those with homomorphic threshold crypto systems [CDN01] and protocols from secure multiparty computation [ST04] to outsource the computations in a blinded way. This leads to the outsource protocols with the following properties:

1. None of the VDs stores the master-key.
2. The master-key remains hidden to any one except for collusions involving a sufficient number of corrupted servers.
3. Secure servers do not obtain information about stored templates as long as not too many servers are corrupted.
4. The new model is vulnerable to a new attack in which the adversary experiments with the secure servers by impersonating a VD. We show that this attack does not compromise the security and privacy in our final protocol.
5. The final protocol leads to our biometrics verification scheme where we show that keys can be renewed by the secure servers without additional interaction with Alice.

REFERENCES

[AC93] R. Ahlswede and I. Csiszár, Common Randomness in Information Theory and Cryptography - Part I: Secret Sharing, *IEEE Transactions on Information Theory*, 39(4):1121–1132, 1993.

[B02] R.M. Bolle, J. Connell, S. Pankanti, N. Ratha, Biometrics 101, Report RC22481, IBM Research, 2002.

[B] J. Bolling, A window to your health, In *Jacksonville Medicine*, 51, Special Issue: Retinal deseases, 2000.

[CDN01] R. Cramer and I. Damgård and J.B. Nielsen, Multiparty Computation from Threshold Homomorphic Encrytpion, In *Advances in Cryptology - Eurocrypt'01*, LNCS 2045, Springer-Verlag, Berlin, 280–300, 2001.

[DRS04] Y. Dodis, L. Reyzin, A. Smith, Fuzzy Extractors: How to generate strong secret keys from biometrics and other noisy data, In *Advances in Cryptology - Eurocrypt'04*, LNCS 3027, 523–540, 2004.

[GT04] J. Goseling and P. Tuyls, Information theoretic approach to privacy protection of biometric templates, *Proceedings of the 2004 International Symposium on Information Theory (ISIT 2004)*, p172, Chicago, 2004.

[JS02] A. Juels, M. Sudan, A Fuzzy Vault Scheme *Proceedings of the 2002 International Symposium on Information Theory (ISIT 2002)*, Lausanne.

[JW99] A. Juels and M. Wattenberg, A fuzzy commitment scheme, *6th ACM Conference on Computer and Communication Security*, p. 28-36, 1999.

[LT03] J.-P. Linnartz and P. Tuyls, New shielding functions to enhance privacy and prevent misuse of biometric templates, *4th International Conference on Audio- and Video-Based Biometric Person Authentication*, 2003.

[MMJ+03] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*, Springer-Verlag New-York 2003.

[MvOV96] A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1996.

[M93] U. Maurer, Secret Key Agreement by Public Discussion from Common Information, *IEEE Transactions on Information Theory*, 39(4):733–742, 1993.

[P65] L. Penrose, Dermatoglyphic topology, *Nature*, 205, (1965) 545-546.

[PK00] T. van de Putte and J. Keuning Biometrical Fingerprint Recognition: Don't get your fingers burned, *IFIP TC8/WG8.8 Fourth working conference on Smart-Card Research and Advance Applications*, Kluwer Academic Publishers (2000), 289-303.

[S99] B. Schneier, Inside risks: The uses and abuses of Biometrics, *Communications of the ACM*, 42, p136, 1999.

[SRS+98] C. Soutar, D. Roberge, S.A. Stojanov, R. Gilroy and B.V.K. Vijaya Kumar, Biometric Encryption-Enrollment and Verification Procedures, Proc. of SPIE, Vol. 3386, 24-35, April 1998.

[ST04] B. Schoenmakers and P. Tuyls, Efficient General Two-Party Computation under the DDH assumption, accepted at Asiacrypt 2004.

[TG04] P. Tuyls and J. Goseling, Capacity and Examples of Template Protecting Biomet-

ric Authentication Systems, *Biometric Authentication Workshop (BioAW 2004)*, LNCS 3087, 158–170, Prague, 2004.

[VTD+03] E. Verbitskiy, P. Tuyls, D. Denteneer, and J.P. Linnartz, Reliable biometric authentication with privacy protection, *The IEEE Benelux Symp. on Inf. Theory*, Veldhoven, The Netherlands, 2003.

[UPP+04] U. Uludag, S. Pankanti, S. Prabhakar, and A.K. Jain, Biometric Cryptosystems: Issues and Challenges, In *Proceedings of the IEEE*, Vol. 92, 6, June 2004.