La biométrie révocable : principes et limites

Estelle Cherrier, Patrick Lacharme, Christophe Rosenberger

La biométrie révocable : principes et limites

Estelle Cherrier, Patrick Lacharme et Christophe Rosenberger
prenom.nom@ensicaen.fr

Résumé

Le développement des systèmes biométriques entraîne des nouvelles menaces et vulnérabilités sur la vie privée des personnes. Cet article présente les techniques de biométrie révocable utilisées pour la protection des données et les limites de tels systèmes.

1 Introduction

La biométrie est une technologie émergente qui propose de nouveaux facteurs d'authentification pour des applications variées. Les schémas actuels sont basés sur de multiples modalités allant de la reconnaissance faciale, les empreintes digitales jusqu'à la biométrie comportementale comme la dynamique de frappe sur un clavier. Le développement important des systèmes biométriques s'accompagne également de plusieurs menaces spécifiques à cette technologie, car les données biométriques sont des données personnelles, non-révocables et donc particulièrement sensibles. Les attaques et les vulnérabilités sur les systèmes biométriques sont nombreuses et peuvent grandement porter atteinte à la vie privée d'un individu.

Les systèmes d'authentification et d'identification biométrique utilisent deux phases distinctes : l'enrôlement et la vérification. Lors de l'enrôlement, l'utilisateur présente une donnée biométrique qui est modélisée puis stockée pour servir de référence. La phase de vérification (*matching*) consiste à comparer la nouvelle donnée capturée, soit avec toutes les références de la base (identification), soit avec la référence de l'utilisateur présumé (authentification). Le modèle de sécurité traditionnel des systèmes biométriques, décrit par Ratha *et al.* [24], comprend notamment les vulnérabilités suivantes : la possibilité de présenter une fausse donnée biométrique au capteur (usurpation d'identité) ; la compromission de la base de données de référence ; la récupération des données biométriques (ou secrètes) qui transitent dans le système ou durant le procédé de comparaison ; la présentation par un attaquant de sa propre donnée. Certaines attaques sont spécifiques à la modalité biométrique considérée : on peut citer le cas de données à trace comme les empreintes digitales, qui ont un statut particulier pour la CNIL [9].

La mise en place de schémas assurant la protection des données biométriques est indispensable et fait l'objet d'une recherche spécifique depuis une dizaine d'années. Parmi les systèmes de protection des données, l'utilisation d'un système d'authentification de type *match on card* stocke la référence biométrique

de l'individu et effectue la comparaison directement sur une carte à puce. La mise en place de ce système n'est toutefois pas toujours possible selon l'application (notamment pour des raisons de temps de calcul) et implique par ailleurs que la carte soit inviolable. Les mécanismes de protection de données sont généralement divisés en deux familles : les mécanismes de chiffrement biométrique et ceux utilisant la biométrie révocable [16, 6, 1, 27]. Le chiffrement classique d'une donnée biométrique ne garantie par sa sécurité dans la mesure où le déchiffrement de la donnée est nécessaire pour effectuer la comparaison avec une capture. Il est également difficile de garantir qu'une donnée biométrique chiffrée ne soit pas décryptée par un attaquant pendant la durée de vie de l'individu. Seule la biométrie révocable est considérée dans cet article.

2 Biométrie révocable

L'expression *biométrie révocable* est définie pour la première fois dans les articles [24] et [4]. Ce concept repose sur une transformation des données biométriques brutes, de telle sorte que les données transformées soient sûres et respectueuses de la vie privée, en accord avec les propriétés détaillées par Maltoni *et al.* [20] :

- *Non-inversibilité :* il ne doit pas être possible de retrouver des informations sur la donnée biométrique originale à partir de sa transformée.
- *Performance :* l'efficacité du système de vérification ne doit pas être détériorée par la transformation.
- *Diversité :* il doit être possible de générer plusieurs données protégées à partir d'une seule donnée brute. Le recoupement de différentes données protégées ne doit pas affecter la protection de la vie privée.
- *Révocabilité :* on doit pouvoir facilement révoquer les données en cas de compromission.

Les techniques de biométrie révocable permettent d'éviter de stocker les données originales : seules les données transformées sont conservées pour la vérification. La propriété de révocabilité est ainsi garantie : si une donnée transformée est compromise, il suffit de changer (les paramètres de) la fonction de transformation. La propriété de diversité est également assurée par le choix de fonctions différentes pour des applications distinctes. En outre, le système de vérification doit être sensible aux variations inter-classe (i.e. pouvoir distinguer deux utilisateurs différents) et à la fois robuste aux variations intra-classe (la donnée biométrique d'un utilisateur varie inévitablement, à cause de conditions de captures différentes, du vieillissement...). Pour cela, les transformations de données biométriques utilisent une donnée ou clé secrète en plus de la donnée biométrique originale. L'enrôlement consiste à calculer la transformée de la donnée de référence à l'aide de la clé, puis à stocker cette donnée transformée. La vérification nécessite le calcul de la transformée de la donnée présentée avec la clé de l'utilisateur, et la comparaison est effectuée entre les données transformées uniquement.

De nouvelles contraintes, spécifiques à la biométrie révocable apparaissent. Le risque de compromission pèse sur la donnée biométrique transformée, mais aussi sur le secret. Par conséquent, ces deux données ne doivent pas être stockées ensemble. Par ailleurs, la compromission de l'une de ces deux données ne doit

pas permettre une usurpation d'identité. De la même façon, l'interception de données transformées (avec différents secrets) correspondant à des applications différentes ne doit pas permettre de remonter au secret, ni à la donnée originale.

Il y a deux types de transformations utilisées pour la biométrie révocable. Le premier type de transformation s'applique directement sur l'image de la donnée biométrique et utilise en général une représentation de taille maximale et fixe (par exemple des attributs de texture calculés sur l'image) [7, 24, 3]. La donnée biométrique brute est ainsi modélisée par un vecteur de réels. Les descripteurs à base de texture (Gabor, LBP...) permettent de générer des vecteurs de taille fixe, contrairement aux minuties dans le cas des empreintes digitales. Le second type de transformation prend en entrée un vecteur binaire ou réel, extrait à partir de la donnée biométrique. Le respect des propriétés de robustesse intra-classe et de sensibilité inter-classe, ainsi que la nature du vecteur de données originales vont avoir un impact sur le choix de la transformation [21]. Les transformations de données biométriques les plus simples sont celles qui sont appliquées aux données binaires, comme l'iriscode. L'iriscode est un vecteur binaire de taille 2048 bits dérivé de l'image d'un iris selon la méthode proposée et régulièrement améliorée par Daugman [10, 11, 12]. La sécurité de la biométrie basée sur l'iris est largement discutée dans [26]. Plusieurs transformations sont proposées, utilisant toutes une donnée secrète pour diversifier et masquer l'iriscode [23, 32]. Si cette donnée est compromise, il est possible soit de retrouver l'iriscode original, soit de construire facilement un autre iriscode dont l'image serait identique à la donnée transformée. La transformation binaire révocable la plus sécurisée est connue sous le nom d'engagement flou, basée sur les codes correcteurs. Néanmoins, cette technique ne peut être appliquée qu'à des données binaires de taille similaire à un iriscode et ne peut donc pas être utilisée efficacement sur n'importe quelle modalité biométrique. La section suivante fait l'objet d'une transformation appliquée sur une telle donnée.

3 BioHachage

Dans cette section, on s'intéresse à un schéma de biométrie révocable très populaire, appelé BioHachage principalement appliqué sur les empreintes digitales. Ce schéma a été proposé en 2003 pour la reconnaissance faciale [13], puis en 2004 pour les empreintes digitales [30]. Lors de la phase d'enrôlement, l'utilisateur présente son empreinte et la clé secrète stockée sur une clé USB, une carte à puce... ou plus généralement un *token*. Des paramètres sont extraits de l'empreinte (par exemple à l'aide d'un banc de filtres de Gabor, [15, 14, 3]) sous forme de FingerCode. La fonction de transformation prend comme entrée ce FingerCode et la clé secrète pour générer un BioCode binaire. Ce BioCode est ensuite stocké dans la base de données. Lors de la phase de vérification, un nouveau BioCode est calculé et comparé au BioCode de référence. La transformation comporte deux étapes : la projection de la donnée biométrique originale par une matrice pseudoaléatoire orthonormale (générée à partir de l'aléa stocké sur le token), suivie d'une quantification en fonction d'un seuil prédéfini, comme présenté figure 1.

Les performances des différents algorithmes de Biohachage étaient initiale-

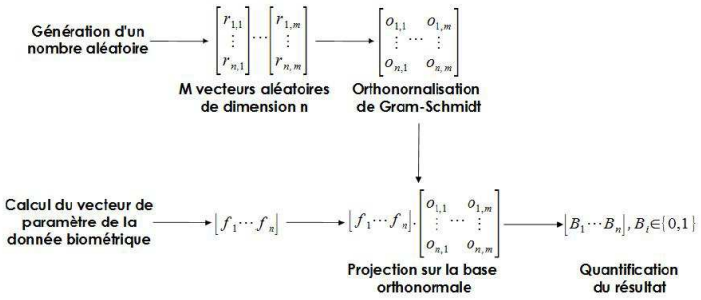

FIGURE 1 – Génération du BioCode

ment annoncées par leurs auteurs comme excellentes, avec un EER affiché de 0%, au lieu des 20% généralement obtenus avec des empreintes digitales. Des études plus précises [17, 25] montrent qu'un tel résultat repose sur la présence de la donnée secrète, et ne peut être obtenu que dans des conditions idéales. Différentes améliorations ont été proposées, comme une approche multimodale [18, 19], l'amélioration de l'extraction des paramètres pour les empreintes digitales [22] et quelques variantes [29, 31]. Il a aussi été proposé de combiner cette technique avec des engagements flous basés sur les codes correcteurs [28, 5].

Des études sur la sécurité de la biométrie révocable ont été développées, avec un nouveau formalisme, les métriques correspondantes et de nouveaux critères de sécurité [21, 2]. Ainsi, la figure 2 présente la probabilité de réussite d'une attaque sur ce schéma révocable, sous plusieurs hypothèses, notamment dans le cas où plusieurs biocodes provenant de diverses applications seraient connus. Ces attaques ont été réalisées sur la base de donnée d'empreinte digitales FVC2002 et présentée dans [2]. Le principe est de prendre plusieurs hypothèses de compromission par un attaquant et de tester l'efficacité de l'attaque. On y distingue clairement l'évolution de l'efficacité de ces attaques sous différentes hypothèses.

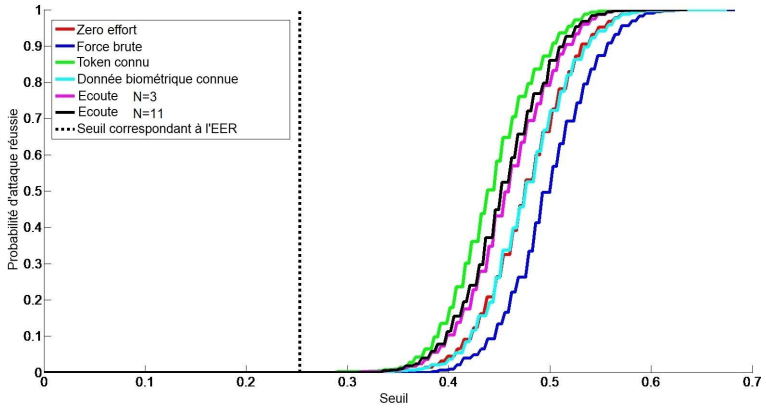

FIGURE 2 – Attaques sous des hypothèses variées

Bien que ce schéma soit a priori basé sur un algorithme inversible, la connais-

sance de la donnée secrète et du BioCode permet d'approximer une donnée biométrique, différente de la donnée originale, mais dont la donnée transformée sera acceptée lors de la phase de vérification. Une usurpation d'identité est ainsi possible, comme cela a été réalisé pour la biométrie faciale et les empreintes digitales [21, 22]. La sécurité de l'algorithme de BioHachage est donc complètement reliée à la protection de la donnée secrète.

4 Conclusion

La biométrie révocable est utilisée pour diversifier et sécuriser les données biométriques, afin de ne pas utiliser directement les données originales et protéger ainsi la vie privée des personnes. Ces schémas dépendent cependant de la modalité biométrique utilisée et doivent nécessairement prendre en compte la sécurité des données et les contraintes de variabilité entre les données. La construction de tels schémas est un important challenge pour la protection de la vie privée et la normalisation de l'évaluation de la robustesse de ces techniques est en cours.

Références

[1] V. Alimi, R. Belguechi, E. Cherrier, P. Lacharme, and C. Rosenberger. *An Overview on Privacy Preserving Biometrics*. Book on Biometrics, InTech, 2011.

[2] R. Belguechi, E. Cherrier, and C. Rosenberger. How to evaluate transformation based cancelable biometric systems ? In *NIST International Biometric Performance Testing Conference (IBPC)*, 2012.

[3] R. Belguechi, B. Hemery, and C. Rosenberger. Authentification révocable pour la vérification basée texture d'empreintes digitales. In *Congrès Francophone en Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, 2010.

[4] R.M. Bolle, J.H. Connel, and N.K. Ratha. Biometric perils and patches. *Pattern Recognition*, 35(12) :2727–2738, 2002.

[5] J. Bringer, H. Chabanne, and B. Kindarji. The best of both worlds : Applying secure sketches to cancelable biometrics. *Sci. Comput. Program.*, 74((1-2)) :43–51, 2008.

[6] A. Cavoukian and A. Stoianov. *Biometric Encryption*. The Encyclopedia of Biometrics, Springer Verlag, 2009.

[7] S. C. Chong, A. B. J. Teoh, and D. C. L. Ngo. High security iris verification system based on random secret integration. *Computer Vision and Image Understanding*, 102(2) :169–177, 2006.

[8] S. C. Chong, A. B. J. Teoh, and D. C. L. Ngo. Iris authentication using privatized advanced correlation filter. In *International Conference in Biometrics (ICB)*, pages 382–388, 2006.

[9] CNIL. Communication de la CNIL relative à la mise en oeuvre de dispositifs de reconnaissance par empreinte digitale avec stockage dans une base de données.

[10] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11) :1148–1161, 1993.

[11] J. Daugman. The importance of being random : Statistical principles of iris recognition. *Pattern Recognition*, 36(2) :279–291, 2003.

[12] J. Daugman. How iris recognition works. *IEEE Transactions on Circiuts and Systems for Video Technology*, 14(1) :21–30, 2004.

[13] A. Goh and C. Ngo. *Computation of Cryptographic Keys from Face Biometrics*, volume 2828 of *LNCS*. Springer, Berlin, 2003.

[14] L. Hong, Y. Wan, and A. Jain. Fingerprint image enhancement : algorithm and performance evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8) :777–789, 1998.

[15] A. Jain, S. Prabhakar, L. Hong, and S. Pankanti. Filterbank-based fingerprint matching. *IEEE Trans. on Image Processing*, 9(5) :846–859, 2000.

[16] A. K. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP J. Advances in Signal Processing*, 8(2) :1–17, 2008.

[17] A. Kong, K.H. Cheung, D. Zhang, M. Kamel, and J. You. An analysis of biohashing and its variants. *Pattern Recognition*, 39, 2005.

[18] A. Lumini and L. Nanni. Empirical tests on biohashing. *NeuroComputing*, 69 :2390–2395, 2006.

[19] A. Lumini and L. Nanni. An improved biohashing for human authentication. *Pattern Recognition*, 40 :1057–1065, 2007.

[20] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer, 2009.

[21] A. Nagar, K. Nandakumar, and A. K. Jain. Biometric template transformation : A security analysis. In *SPIE, Electronic Imaging, Media Forensics and Security XII*, 2010.

[22] L. Nanni and A. Lumini. Local binary patterns for a hybrid fingerprint matcher. *Pattern Recognition*, 41(11) :3461 – 3466, 2008.

[23] O. Ouda, N. Tsumura, and T. Nakaguchi. Tokenless cancelable biometrics scheme for protecting iris codes. In *20th International Conference on Pattern Recognition (ICPR'10)*, pages 882–885, 2010.

[24] N. Ratha, J. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics based authentication systems. *IBM Systems*, 40(3) :614–634, 2001.

[25] C. Rathgeb and A. Uhl. Two-factor authentication or how to potentially counterfeit experimental results in biometric systems. In *ICIAR*, pages 296–305, 2010.

[26] C. Rathgeb and A. Uhl. *The State-of-the-Art in Iris Biometric Cryptosystems*. InTech, 2011.

[27] C. Rathgeb and A. Uhl. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP J. on Information Security*, 3, 2011.

[28] A.B.J. Teoh, B. Jin, T. Connie, D. Ngo, and C. Ling. Remarks on biohash and its mathematical foundation. *Information Processing Letters*, 100(4) :145–150, 2006.

[29] A.B.J. Teoh, Y. Kuanb, and S. Leea. Cancellable biometrics and annotations on biohash. *Pattern recognition*, 41 :2034–2044, 2008.

[30] A.B.J. Teoh, D. Ngo, and A. Goh. Biohashing : two factor authentication featuring fingerprint data and tokenised random number. *Pattern recognition*, 40, 2004.

[31] A.B.J. Teoh, K.-A. Toh, and W. Yip. 2^n discretisation of biophasor in cancellable biometrics. In *Advances in Biometrics*, volume 4642 of *LNCS*, pages 435–444. Springer Berlin / Heidelberg, 2007.

[32] J. Zuo, N. K. Ratha, and J. H. Connell. Cancelable iris biometric. In *Conference on Pattern Recognition (ICPR 2008)*, pages 1–4, 2008.