Fuzzy Commitment Schemeを用いたバイオメトリック暗号における テンプレートの安全性に関する一考察

披田野 清良† 市野 将嗣‡ 大木 哲史† 高橋 健太* 小松 尚久†

†早稲田大学理工学術院
169-8555 東京都新宿区大久保 3-4-1
{hidano, ohki, komatsu}@kom.comm.waseda.ac.jp

‡電気通信大学大学院情報理工学研究科
182-8585 東京都調布市調布ヶ丘 1-5-1
ichino@inf.uec.ac.jp

*株式会社日立製作所横浜研究所
244-0817 神奈川県横浜市戸塚吉田町 292
東京大学大学院情報理工学系研究科
113-8656 東京都文京区本郷 7-3-1
kenta.takahashi.bw@hitachi.com

あらまし 近年，生体情報を秘匿して認証を行うテンプレート保護型生体認証が注目されている．しかしながら，それらの安全性に関する議論では，生体情報間の相関性により，生体情報の推定が容易となる可能性については必ずしも十分に検討されていない．そこで，本稿では，Fuzzy Commitment Scheme を用いたバイオメトリック暗号に着目し，まず，Renyi エントロピーを用いた情報量評価手法に基づき，一例として指紋情報の情報量を定量的に評価することにより，生体情報間に強い相関性があることを明らかにする．次いで，上記の相関性を利用したなりすましに関する脅威を示し，それらの脅威に対する安全性について考察する．

A Study on Security of Template in Biometric Cryptosystem Using Fuzzy Commitment Scheme

Seira HIDANO† Masatsugu ICHINO‡ Tetsushi OHKI†
Kenta TAKAHASHI* Naohisa KOMATSU†

†Faculty of Science and Engineering, Waseda University
3-4-1 Ohkubo, Shinjuku-ku, Tokyo, 169-8555, JAPAN
{hidano, ohki, komatsu}@kom.comm.waseda.ac.jp

‡Graduate School of Informatics and Engineering, The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585, JAPAN
ichino@inf.uec.ac.jp

* Yokohama Research Laboratory, Hitachi, Ltd.
292 Yoshida-cho, Totsuka-ku, Yokohama-shi, Kanagawa-ken, 244-0817, JAPAN
Graduate School of Information Science and Technology, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, JAPAN
kenta.takahashi.bw@hitachi.com

Abstract Template protection-based biometric authentication attracted attention in the past decade. However, in the discussion on the security of these systems, biometric information content is assumed to be sufficiently large and real conditions of biometric samples have not yet been reflected. Thus, it is demonstrated in this paper that there is any correlation between biometric samples by evaluating fingerprint information content on the basis of an evaluation method using a Renyi entropy. Additionally, spoofing attacks taking advantage of the correlation are explained and the security against these attacks is discussed.

1 はじめに

生体認証は，記憶，所持の煩わしさから解放されるなどの利便性がある一方，ユーザならびに環境条件，運用条件，生体情報，認証装置などの様々な構成要素において特有の脆弱性が存在する．特に，生体認証においてシステムに保管されている生体情報（以下，テンプレート）は，個人性を多く含む機微情報であり，変更することができないため，情報漏洩に対するリスクが非常に大きい．

このため，近年，生体情報を解読不可能な状態に変換して認証を行うテンプレート保護型生体認証が注目されている．代表的な研究事例としては，ユーザが提示する生体情報とテンプレートから一意の秘密鍵を生成するバイオメトリック暗号 [1, 2] や，幾何変形もしくは相関性のあるハッシュ関数などを用いてテンプレートを非可逆に変換するキャンセラブルバイオメトリクス [3] などがあげられる．また，保護テンプレートから元の生体情報を復元する困難さはテンプレート保護技術の安全性評価項目の一つであり [4]，当該安全性を情報理論の概念に基づき理論的に考察する試みもある [5, 6, 7]．しかしながら，それらの議論では，生体情報の情報量が十分に大きいことを前提としており，生体情報間の相関性により，生体情報の推定が容易となる可能性については必ずしも十分に検討されていない．

そこで，本稿では，テンプレートの安全性のみならず，秘密鍵の秘匿性にも優れており，Challenge Handshake Authentication Protocol（CHAP）などの既存の認証プロトコルとの親和性が高く実用的な手法であることから，Fuzzy Commitment Scheme（FCS）を用いたバイオメトリック暗号に着目し，生体情報間の相関性を考慮してテンプレートの安全性について検討する．具体的には，まず，筆者らが提案する Renyi エントロピーを用いた生体情報の情報量評価手法 [8] に基づき，一例として指紋情報の情報量を定量的に評価することにより，生体情報間に何らかの相関性があることを明らかにする．次いで，生体情報間の相関性を利用したなりすましに関する脅威を示し，それらの脅威に対する安全性を理論的に考察すると共に，Renyi エントロピー評価と同様に指紋認証を一例として，シミュレーション結果を交えて定量的に評価する．

2 Fuzzy Commitment Scheme を用いたバイオメトリック暗号

Fuzzy Commitment Scheme（FCS）は，1999 年に Juels らにより提案された誤り訂正符号を用いた暗号方式の一種である [1]．図 1 に，サーバ/クライアントモデルにおける FCS を用いたバイオメトリック暗号の概要を示す．

登録過程

1. クライアントは，ユーザが提示した生体に関する原情報からビット列 $b \in \mathcal{B} \subseteq \{0,1\}^n$ を抽出する．ただし，$|\mathcal{B}| \leq 2^n$ とする．$|\cdot|$ は集合の要素数を示し，以下，b を生体情報と呼ぶ．

2. クライアントは，ランダムに選択した秘密鍵 $s \in \{0,1\}^k$ から誤り訂正符号化により符号語 $c \in \mathcal{C}$ を生成し，b と c の排他的論理和によりコミットメント $z = b \oplus c$ を作成する．ただし，本稿では，$\mathcal{C}$ は，符号長 n，情報記号数 k，最小距離 d_{min} の (n, k, d_{min})-線形符号とする．このとき，$|\mathcal{C}| = 2^k$ となり，$t = (d_{min} - 1)/2$ 個以下のビット誤りを訂正できる．

3. クライアントは，c のハッシュ値 $h(c)$ を計算し，z と $h(c)$ の組 $(z, h(c))$ を認証サーバに送信する．

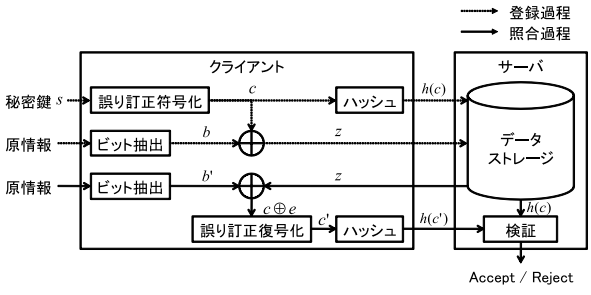

図 1: FCS を用いたバイオメトリック暗号

4. 認証サーバは，$(z, h(c))$ をストレージに保管する．

照合過程

1. クライアントは，登録時と同様に，ユーザが提示した生体に関する原情報から b と同形式の生体情報 $b' \in \mathcal{B}$ を抽出し，当該ユーザの z を認証サーバから取り寄せる．
2. クライアントは，b' と z の排他的論理和を計算し，$c \oplus e = b' \oplus z$ から誤り訂正復号化により c' を得る．ただし，$e = b \oplus b'$ とし，$\| e \| \leq t$ のとき，c' は c と一致する．$\| \cdot \|$ はハミング重みを示す．
3. クライアントは，c' のハッシュ値 $h(c')$ を計算し，$h(c')$ を認証サーバに送信する．
4. 認証サーバは，$h(c')$ と $h(c)$ を比較し，ユーザの正当性を検証する．

3 指紋情報の情報量評価

本章では，Renyi エントロピーを用いた生体情報の情報量評価手法 [8] に基づき，指紋情報の情報量を定量的に評価する．

3.1 生体情報の情報量評価手法

生体情報 $b \in \mathcal{B}$ を取り得る値とする確率変数 B の 2 次の Renyi エントロピー（以下，Renyi エントロピー）$H_2(B)$ は，B の確率関数 $p_B(b)$ を用いて，次式で表せる．

$$H_2(B) = -\log_2 \sum_{b \in \mathcal{B}} p_B(b)^2 \tag{1}$$

(1) 式の $\sum_{b \in \mathcal{B}} p_B(b)^2$ は $\mathcal{B}$ 上の 2 つの確率変数が同一の値を取る確率を示しており，このとき，2 つの生体情報間の距離は 0 となる．このため，$H_2(B)$ は，生体情報間の距離 $d \in \mathbb{R}$ を取り得る値とする確率変数 D の確率関数 $p_D(d)$ を用いて，次式で表せる．

$$H_2(B) = -\log_2 p_D(0) \tag{2}$$

$p_D(0)$ は $p_D(d)$ が既知であれば容易に導出できる．Daugman の虹彩認証モデル [9] のように $p_D(d)$ の形状が十分に検討された生体情報であれば，$p_D(d)$ は他人間照合実験を通して得られる生体情報間の距離のサンプルを用いてパラメトリックに推定する．一方，$p_D(d)$ の形状が未知でモデル化が困難な場合は，分布の形状を仮定しないノンパラメトリックな手法を用いて推定する．

3.2 FCS における生体情報の Renyi エントロピー

2 章で示したように，FCS を用いたバイオメトリック暗号では，生体情報 b は符号長 n のビット列 $\mathcal{B} = \{0,1\}^n$ で記述され，2 つの生体情報 $b, b' \in \mathcal{B}$ の間の距離 d は次式で表せる．

$$d = \frac{\| b \oplus b' \|}{n} \tag{3}$$

このとき，d を取り得る値とする確率変数 D の確率関数 $p_D(d)$ は，b の記述形式および d の定義が Daugman の虹彩コードを用いた虹彩認証モデルと同様であることから，次式に示す二項分布 $Bi(\theta, \hat{n})$ でモデル化できると仮定する．

$$p_D(d) = \frac{\hat{n}!}{(\hat{n}d)!(\hat{n}(1-d))!} \theta^{\hat{n}(1-d)} (1-\theta)^{\hat{n}d} \tag{4}$$

D の平均 $E(D)$ および分散 $V(D)$ はそれぞれ次のように表せる．

$$E(D) = 1 - \theta \tag{5}$$

$$V(D) = \frac{\theta(1-\theta)}{\hat{n}} \tag{6}$$

(4) 式より，$p_D(0) = \theta^{\hat{n}}$ となり，$\mathcal{B}$ 上の確率変数 B の Renyi エントロピー $H_2(B)$ は次式で表せる．

$$H_2(B) = -\log_2 \theta^{\hat{n}} \tag{7}$$

$p_D(d)$ を二項分布でモデル化した場合，θ は b の各ビットの一致確率，$\hat{n}$ は b のビット間に何らかの相関がある場合の見かけ上の有効ビット数と解釈できる．また，それらの値は，生体情報のサンプルを用いた他人間照合実験を通して得られる距離のサンプルの平均および分散と (5) 式，(6) 式より推定する．

3.3 情報量評価実験

FCS を用いたバイオメトリック暗号を指紋認証に適用し，指紋画像のデータベースを用いて他人間照合実験を行い，指紋情報の情報量を定量的に評価した結果について述べる．

本実験では，Tuyls らの指紋ビット列を用いたバイオメトリック暗号 [10] を評価対象とし，指紋画像データベースとして FVC2002 DB1 のセット A に収録されている異なる 100 指から 8 枚ずつ取得した計 800 枚の画像を使用した．ただし，それぞれの指について 6 枚を登録用，残りの 2 枚を照合用とし，符号長 $n = 127$ の指紋ビット列を生成して照合を行った．

その結果，指紋情報間の距離の平均は 0.499，分散は 0.00684 であった．(5) 式および (6) 式より，各ビットの一致確率 θ の推定値は 0.501，有効ビット数 $\hat{n}$ の推定値は 37 となり，(7) 式より，指紋情報 $b \in \mathcal{B}$ を取り得る値とする確率変数 B の Renyi エントロピー $H_2(B)$ は 36 bits であった．$\mathcal{B} = \{0,1\}^{127}$ かつ B が一様分布に従うとき，$H_2(B)$ は理想的に 127 bits となる．しかしながら，$H_2(B)$ の実験値はこの値を大きく下回ることから，指紋情報間には何らかの相関性があると言える．これらの相関性を利用した脅威については 4 章に示す．

4 FCS に対する攻撃

本章では，強い条件を仮定せずに容易に実現可能な生体情報間の相関性を利用したなりすましに関する脅威を示す．まず，通常運用時の脅威として，生体情報間の相関性のみを利用した Biometric Dictionary Attack（BDA）を示す．ただし，通常運用時は，符号語空間の大きさのみを利用した Exhaustive Codeword Search Attack（ECSA）も容易に実現可能であり，ECSA の攻撃成功確率が BDA のそれを上回る場合，攻撃者は ECSA によりなりすましを試みる可能性がある．このため，次いで ECSA の攻撃手順を示す．最後に，コミットメントが漏洩した際の脅威として，生体情報間の相関性と符号語空間の大きさを同時に考慮した Decodable Biometric Dictionary Attack（DBDA）について述べる．

4.1 Biometric Dictionary Attack（BDA）

BDA の攻撃手順を以下に示す．

1. システムで利用されているモダリティの生体情報データベース $DB = \{b_1, \ldots, b_N\}$ を用意し，DB からランダムに 1 つを攻撃用生体情報 b^* として選択する．
2. 2 章で示した照合過程の手順 1 において，攻撃対象ユーザになりすまし，b^* をクライアントに入力する．
3. 照合過程の手順 4 において，認証サーバが Accept を返せば攻撃成功とする．

4.2 Exhaustive Codeword Search Attack（ECSA）

ECSA の攻撃手順を以下に示す．ただし，攻撃者は誤り訂正符号に関する各パラメータと生成多項式を知っているものとする．

1. 2^k 個の符号語空間 $\mathcal{C}$ からランダムに 1 つを攻撃用符号語 c^* として選択する．
2. 照合過程の手順 3 において，攻撃対象ユーザになりすまし，c^* をクライアントに入力する．
3. 照合過程の手順 4 において，認証サーバが Accept を返せば攻撃成功とする．

4.3 Decodable Biometric Dictionary Attack（DBDA）

あるユーザのコミットメント $z = b \oplus c$ が漏洩した際の DBDA の攻撃手順を以下に示す．ただし，本稿では，攻撃時に当該ユーザのテンプレートの更新は行われていないものとする．

1. BDA と同様にシステムで利用されているモダリティの生体情報データベース $DB = \{b_1, \ldots, b_N\}$ を用意する．

2. $b_i \in DB$ と取得した z との排他的論理和を計算し，$b_i \oplus z$ に対して誤り訂正復号化処理を施す．ただし，このとき，$b_i \oplus z$ は何らかの符号語に復元される場合と復号化に失敗して何も復元されない場合がある．

3. 手順 2 において，$b_i \oplus z$ が何らかの符号語に復元された場合，b_i を新たな攻撃用生体情報データベース $\overline{DB}$ に加える．手順 2 および手順 3 はすべての b_i について行う．

4. $\overline{DB}$ からランダムに 1 つを攻撃用生体情報 b^* として選択して，照合過程の手順 1 において，z が盗まれたユーザになりすまし，b^* をクライアントに入力する．

5. 照合過程の手順 4 において，認証サーバが Accept を返せば攻撃成功とする．

5 セキュリティ分析

4 章で示したそれぞれの脅威に対する安全性について，5.1 節および 5.2 節において理論的に考察し，5.3 節においてシミュレーション結果を交えて定量的に評価する．

5.1 通常運用時の攻撃成功確率

まず，BDA の攻撃成功確率は，生体認証の標準的な精度評価尺度の一つである False Accept Rate（FAR）と一致する．FAR は，$\{0,1\}^n$ からランダムに選択した語 x と生体情報空間 $\mathcal{B} \subseteq \{0,1\}^n$ の大きさ $|\mathcal{B}|$ を用いて，次式で表せる．

$$FAR = \frac{|\mathcal{B}_t(b)|}{2^n \cdot P(x \in \mathcal{B})} \quad (8)$$

$$= \frac{|\mathcal{B}_t(b)|}{|\mathcal{B}|} \quad (9)$$

ただし，x および生体情報 $b \in \mathcal{B}$ の確率関数はそれぞれ一様分布と仮定し，$\mathcal{B}_t(b)$ は b を中心とする半径 t の超球の内側にある生体情報の集合，すなわち，$\mathcal{B}_t(b) = \{b' | \parallel b \oplus b' \parallel \leq t, b' \in \mathcal{B}\}$ とする．

次に，ECSA の攻撃成功確率 P_{ECSA} は，符号語空間 $\mathcal{C}$ の大きさ $|\mathcal{C}| = 2^k$ を用いて，次式で表せる．

$$P_{ECSA} = \frac{1}{2^n \cdot P(x \in \mathcal{C})} \quad (10)$$

$$= \frac{1}{|\mathcal{C}|} = \frac{1}{2^k} \quad (11)$$

以上より，通常運用時における攻撃成功確率 Successful Attack Probability（SAP）は次式で表せる．

$$SAP = \max\{FAR, P_{ECSA}\} \quad (12)$$

ただし，$\max\{i,j\}$ は i と j のうち大きい値を返す．

5.2 コミットメント漏洩時の攻撃成功確率

DBDA の攻撃成功確率，すなわち，あるユーザのコミットメント z が漏洩した際の攻撃成功確率 $\overline{SAP}$ は，次式で表せる．

$$\overline{SAP} = \frac{|\mathcal{B}_t(b)|}{2^n \cdot P(x \oplus z \in \bigcup_{c \in \mathcal{C}} \mathcal{C}_t(c), x \in \mathcal{B})} \quad (13)$$

$$\approx \frac{|\mathcal{B}_t(b)|}{2^n \cdot P(x \in \bigcup_{c \in \mathcal{C}} \mathcal{C}_t(c) \cap \mathcal{B})} \quad (14)$$

$$= \frac{|\mathcal{B}_t(b)|}{2^n \cdot P(x \in \bigcup_{c \in \mathcal{C}} \mathcal{C}_t(c)) P(x \in \mathcal{B})} \quad (15)$$

$$= \frac{|\mathcal{B}_t(b)|}{2^n \cdot |\mathcal{C}_t(c)| \cdot P(x \in \mathcal{C}) P(x \in \mathcal{B})} \quad (16)$$

ただし，$\mathcal{C}_t(c)$ はある符号語 $c \in \mathcal{C}$ について誤り訂正可能な語の集合，すなわち，$\mathcal{C}_t(c) = \{w | \parallel c \oplus w \parallel \leq t, w \in \{0,1\}^n\}$ とする．また，(14) 式は，線形符号の性質により，$\mathcal{B}$ と $\mathcal{B}$ を z だけ平行移動した 2 つの空間において，何らかの符号語に誤り訂正可能な語の空間の占める割合がほぼ等しくなることを仮定している [7]．ここで，$|\mathcal{B}_t(b)| = |\mathcal{C}_t(c)|$ を仮定すると，(8) 式，(10) 式より，$\overline{SAP}$ は次式で表せる．

$$\overline{SAP} = FAR \cdot \frac{1}{|\mathcal{C}_t(c)| \cdot P(x \in \mathcal{C})} \quad (17)$$

$$= P_{ECSA} \cdot \frac{1}{P(x \in \mathcal{B})} \quad (18)$$

表 1: 攻撃成功確率

(n, k, d_{min})	(127,8,57)	(127,15,55)	(127,22,47)
FRR	0.0214	0.0267	0.0481
FAR	0.00264	0.00232	0.00124
P_{ECSA}	0.00391	3.05×10^{-5}	2.38×10^{-7}
SAP	0.00391	0.00232	0.00124
$\overline{SAP}$	0.228	0.216	0.135

以上より，$\overline{SAP} \geq SAP$ となり，コミットメント漏洩時の安全性は，通常運用時に比べて低下すると考えられる．

5.3 安全性評価実験

FCS を用いたバイオメトリック暗号を指紋認証に適用し，4 章で示した脅威に対する安全性を定量的に評価した結果について述べる．

本実験では，3.3 節の情報量評価実験と同様に，Tuyls らの指紋ビット列を用いたバイオメトリック暗号 [10] を評価対象とする．ただし，(n, k, d_{min})-線形符号として BCH 符号を用いた．

BCH 符号に関する各パラメータを変化させたときの False Reject Rate（FRR）および FAR，P_{ECSA}，SAP，$\overline{SAP}$ を表 1 に示す．ここで，$(n, k, d_{min}) = (127, 22, 47)$ のときに，指紋情報間に何の相関性もなく，指紋情報 b の確率関数が $\{0,1\}^{127}$ 上の一様分布であると仮定した場合，(9) 式より，$FAR = \sum_{i=0}^{23} {}_{127}\mathrm{C}_i / 2^{127} = 8.48 \times 10^{-14}$ となる．しかしながら，表 1 において，$(n, k, d_{min}) = (127, 22, 47)$ のときの FAR の値に注目すると，上記の値とは大きく異なるため，指紋情報間の相関性により安全性が著しく低下していることが分かる．また，どのパラメータ値においても $\overline{SAP}$ は SAP を大きく上回り，5.1 節および 5.2 節で示したように，コミットメント漏洩時の安全性は通常運用時に比べて著しく低下すると言える．

6 まとめと今後の課題

本稿では，まず，Fuzzy Commitment Scheme（FCS）を用いたバイオメトリック暗号に着目し，一例として指紋情報の Renyi エントロピーを定量的に評価することにより，生体情報間には何らかの相関性があることを明らかにした．そして，生体情報間の相関性を利用した脅威に対する安全性を理論的に考察すると共に，シミュレーション結果を交えて定量的に評価した．その結果，コミットメント漏洩時の安全性は通常運用時に比べて著しく低下するという知見を得た．今後は，5.1 節および 5.2 節の議論において生体情報が一様分布に従わない場合の安全性について理論的に考察する．

参考文献

[1] A. Juels, and M. Wattenberg, “A fuzzy commitment scheme,” Proc. 6th ACM Conference on Computer and Communications Security, pp.28-36, 1999.

[2] A. Juels, and M. Sudan, “A fuzzy vault schem,” Proc. IEEE International Symposium on Information Theory (ISIT 2002), p.408, 2002.

[3] N.K. Ratha, J.H. Connell, and R.M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems,” IBM Systems Journal, vol.40, no.3, pp.614-634, 2001.

[4] A.K. Jain, K. Nandakumar, and A. Nagar, “Biometric template security,” EURASIP J. Advances in Signal Processing, 2008.

[5] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,” SIAM J. Computing, vol.38, no.1, pp.97-139, 2008.

[6] K. Simoens, P. Tuyls, and B. Preneel, “Privacy weaknesses in biometric sketches,” Proc. 2009 IEEE Symposium on Security and Privacy, pp.188-203, 2009.

[7] Y. Wang, S. Rane, S.C. Draper, and P. Ishwar, “An information-theoretic analysis of revocability and reusability in secure biometrics,” Proc. 2011 Information Theory and Applications Workshop (ITA2011), pp.1-10, 2011.

[8] 披田野清良，赤尾直彦，小松尚久，高橋健太，“Renyi エントロピーを用いた虹彩情報の情報量評価手法,” 情報処理学会論文誌，vol.52，no.9，2011.

[9] J. Daugman, “The importance of being random: statistical principles of iris recognition,” Pattern Recognition, vol.36, no.2, pp.279-291, 2003.

[10] P. Tuyls, A.H.M. Akkermans, T.A.M. Kevenaar, G.J. Schrijen, A.M. Bazen, and R.N.J. Veldhuis, “Practical biometric authentication with template protection,” Proc. 5th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA2005), pp.436-446, 2005.