多項式の類似度を利用した非対称生体認証

菊池浩明 †　　尾形 わかは ††　　西垣　正勝 †††

† 東海大学大学院工学研究科，259-1292 神奈川県平塚市北金目 1117，kikn@tokai.ac.jp
†† 東京工業大学，183-8512 東京都目黒区大岡山 2-12-1，
††† 静岡大学創造科学技術大学院，432-8011 静岡県浜松市城北 3-5-1，

あらまし　生体情報における特徴量の変動に対して，多項式の根へ特徴量を符号化することで耐性のある生体認証プロトコルを提案する．提案方式は，サーバへ登録した生体情報のサーバとユーザに対する秘匿性，類似した生体情報を有するユーザだけが認証出来ることを保証する．指紋におけるマニューシャにおいて，幾何学的ハッシュを導入して，提案方式を実現できることを示し，実験結果について報告する．

Asymmetric Biometric Authentication Based on Similarity Defined of Intersecton of Two Sets

Hiroaki Kikuchi†　　Wakaha Ogata††　　Masakatsu Nishigaki†††

†Graduate School of Engineering, Tokai University, 1117, Kitakaname, Hiratsuka, Kanagawa, Japan, 1117,
††Tokyo Institute of Technology, 2-12-1, O-okayama, Meguroku, Tokyo, Japan, 152-8550
†††Graduate School of Science and Technology, Shizuoka University,
3-5-1, Johoku, Hamamatsu, Shizuoka, Japan, 432-8011

Abstract　The paper proposes schemes addressing the issue of unstable feature of biometrics by using a new encoding algorithm to embed feature into encrypted polynomial as its roots. The scheme ensures secrecy of biometric information, and the authenticity of legitimate uses who own similar feature to that stored in enrolment. The experimental results using the geometrical hashing fingerprint are given.

1　はじめに

生体情報を本人認証の鍵としよう．この野望にはいくつもの落とし穴があった．一つには指紋をはじめとする我々の生体情報の持つ本質的な不安定さがあり，それに計測時の誤差が加わり特徴量は予測のつかない変動を引き起こす．もう一つは，認証者と検証者の間の対称性である．大きな変動が避けられない生体情報を登録情報と比較するために，検証者は利用者と同じ秘密情報を有する必要があった．

この問題に多くの試みがある．Juel らは誤り訂正符号を導入して，不完全な生体情報からでも秘密を復元できる Fuzzy Vault を提案した [4]. Uludag らは，しきい値付き秘密分散を用いて指紋認証に適用したが，n 個のマニューシャから正しい t 個を選ぶ組み合わせに比例する計算量，$\binom{n}{t} = O(n^t/t)$ がかかる [5]．ヘルパー情報の元で生体情報を変換して登録するキャンセラブル生体認証の研究も多く行われている [6]．その代償として，ヘルパー情報を安全なデバイスに格納して認証時に持ち歩かなくてはならない．[6] に整理されている．結局のところ，表 1 に示す多くの要求条件を満たさなくてはならないと

表 1: 非対称生体認証スキーム $\mathcal{A}$ の要請

（完全性）正しい認証者ならば $\mathcal{A}$ で証明できる．
（健全性）$\mathcal{A}$ で証明できるならば認証者は正しい．
（ゼロ知識性）検証者は，登録情報と証明から一切の情報を得られない．
(あいまい性）認証者の秘密情報の変動を許す．
（非対称性）検証者は認証者の秘密情報を持たない．
（失効可能性）登録情報の更新が出来る．
（記憶装置不要性）認証者は生体情報以外の秘密情報を持たない．登録時の情報へのアクセスもしない．
（効率性）証明は特徴量のサイズ n と変動の大きさ t に対して効率的である．

ころにこの研究の難しさがある．

そこで，本研究では，**多項式の重根判定問題**を応用して，生体情報以外の秘密情報を持たない認証者が自分が登録した生体情報に近い情報を持っていることを証明するプロトコルを提案する．提案方式は，

1. 検証者には認証結果以外の生体情報が漏れない，
2. 利用者は生体情報以外の秘密情報を必要としない，
3. 利用者さえも登録情報を見ることなく，認証時の生体情報との差以外は漏れない，
4. 特徴量のサイズ n に対して，$O(n^2)$ の計算量で効率的に，

リモート認証を実現する．

提案方式の実現可能性を示すため，Lee らによって提案された幾何学的ハッシュ関数 [2] を導入して，提案方式の有効性を実験的に検証する．

2 準備

2.1 準同型性を満すコミットメント [1]

N を誰もその素因数を知らない大きな合成数とし，g, h を Z_N の要素とする．誰も h の離散対数 $\log_g h$ を知らないとする．この時，$E(m,r) = g^m h^r \bmod N$ を m のコミットメントとする．ただし，r は十分大きい乱数とする．このコミットメントは，加法準同型性 $E(m,r) \times E(m',r') = E(m+m', r+r')$, $E(m,r)^x = E(mx, rx)$ を満たす．

2.2 重根を持つ多項式の性質

t 個の重根を持っている n 次の多項式 f,

$$f(x) = (x-\alpha_1)^2 \cdots (x-\alpha_t)^2 (x-\alpha_{t+1}) \cdots (x-\alpha_n) \tag{1}$$

を考えよう．f は n 次の多項式であるので，少なくとも $n+1$ 個の点 x_i における評価値 $y_i = f(x_i)$ があれば一意に決まる．しかし，式 (1) における係数 α の数が高々 $n-t$ 個しかないことに着目すると，$n-t+1$ 個の評価値 y_i からでも，f は一意に決まる．従って，ラグランジェの補間式に基づいて，次の性質が言える．

命題 2.1 互いに異なる $X = \{x_1, \ldots, x_{n-t+1}\}$ についての f 評価値 $y_1, \ldots, y_{n-t+1}$ が与えられたとき，任意の点 β における評価値 $f(\beta)$ は

$$f(\beta) = \sum_{i=1}^{n-t+1} \Lambda(i, X, \beta) f(x_i) \tag{2}$$

で与えられる．ただし，

$$\Lambda(i, X, \beta) = \prod_{j \in X, j \neq i, j \leq t} \frac{(\beta - x_j)^2}{(x_i - x_j)^2} \prod_{j \in X, j \neq i, j > t} \frac{\beta - x_j}{x_i - x_j}$$

とする．

(証明) 式 (2) は n 次のある多項式を表している．$i = 1, \ldots, n-t+1$ について，$\Lambda(i, X, \beta) = 1$ if $\beta = x_i$, $\Lambda(i, X, \beta) = 0$ otherwise であるので，式 (2) は常に $f(x_i) = y_i$ を満たす．よって，式 (2) は $\alpha_1, \ldots, \alpha_{n-t+1}$ 個の線形式の解になっている．(証明終)

2.3 幾何学的ハッシュ関数

指紋には絶対的な基準点が定められないため，登録指紋と認証時の指紋を照合するには位置あわせをする必要がある．この問題に対して，Leeらは, 幾何学的ハッシュ関数 (Geometric Hashing) と呼ぶ次のような位置あわせ不要の方式を提案している [2, 3].

登録マニューシャの集合を $A = \{a_1, \ldots, a_n\}$ とする．ここで, $m_i = (x_i, y_i, \theta_i, t_i)$ は座標 x_i, y_i, 方向 θ_i，種類 t_i（分岐点と端点）を持つ A の要素とする．A の中心からコア（basis) を一つ選び，それを a_1 とする．A の全ての要素 a_i を a_1 を原点とする相対座標 に線形変換する，すなわち，

$$\begin{pmatrix} x_i' \\ y_i \end{pmatrix} = \begin{pmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{pmatrix} \begin{pmatrix} x_i - x_1 \\ y_i - y_1 \end{pmatrix}$$

ここで，$\delta = \theta_i - \theta_1$ である．これを，中心から近い順で k 個のコアについて行い，それを $A_1, \ldots, A_k$ とする．認証時のマニューシャB についても，同様の処理を行い，$B_1, \ldots, B_k$ 個の集合を得る．これらの集合間の最大の積集合の大きさで，A と B の類似度を求める．すなわち，

$$\mathsf{sim} = \max_{1 \le i,j \le k} |\, A_i \cap B_j \,|$$

とする．更に，読み取り時に生じる誤差や生体の変動は避けられないので，適切な定数 q で量子化して，$\mathrm{int}(x_i/q) = \mathrm{int}(x_j/q)$ ならば $x_i = x_j$ とする．

異なるコアについて相対化したマニューシャはほとんど一致することはなく，k についてバースデーパラドックスの原理で $O(k^2)$ のオーダーで一致が起こりやすいことを利用している．

3 提案方式

3.1 概要

登録時の生体情報を $A = \{a_1, \ldots, a_n\}$, 認証時を $B = \{b_1, \ldots, b_n\}$ とする．A と B の要素を根として持つ多項式をそれぞれ $f_A(x)$, $f_B(x)$ とする．すなわち，

$$f_A(x) = (x - a_1) \cdots (x - a_n)$$

とする．この多項式の積で与えられる $2n$ 次の多項式を $F(x) = f_A(x) f_B(x)$ とする．A と B の積集合が十分に大きいとすると，その交わりの数を $t = |A \cap B|$ とおいて，F は t 個の重根を持つ多項式となる．ゆえに，$2n - t + 1$ 個の互いに異なる点からその F の値は一意に決まる．認証者は A を知らなくても，F の重根が与えられた閾値より多いことを証明すればよい．

多項式の重根の数を証明するには，互いに異なる $2n - t + 1$ 個の値についての評価値を 2 組用いて，それぞれから補間した F の評価値が等しいことを示せばよい．ラグランジェの補間法に基づく式 (2) を用いれば，加法準同型性を満たしたコミットメント E で $F(x_i)$ を秘匿したまま，任意の点 β についての評価値のコミットメント $E(F(\beta))$ を計算することが出来る．検証者には，登録された生体情報も認証時に送られる生体情報も秘匿されたままで検証が可能である．

しかし，どの $2n - t + 1$ 個の組からの値も収束すれば，B と A との交わりが少なくとも t 個以下であることは言えるが，これは十分ではない．不正者が，$f_B'(x) = (x - b)^n$ のような根を多重にする多項式を送るかもしれないからである．これを防止するために，認証者は更に，多項式 $f_B(x)$ に互いに異なる n 個の根があることを証明する必要がある．

3.2 提案方式 1

登録フェーズ

登録時と認証時の生体情報をそれぞれ，（マニューシャの）集合 $A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_n\}$ とする．A と B の要素は少なくとも t 個重複していると仮定する．

ユーザは A を根として持つ多項式 $f_A(x) = f_{A_0} + f_{A_1} x + \cdots f_{A_n} x^n$ とランダムな n 次多項式 $r_A(x) = r_0 + r_1 x + \cdots r_n x^n$ を作り，

$$A_i = E(f_{A_i}, r_i) = g^{f_{A_i}} h^{r_i}$$

で与えられるコミットメント $A_0, \ldots, A_n$ をサーバに登録する．

認証フェーズ

1. $n_* = 2n - t + 2$ とする．サーバはランダムな n_* 個の $x_1, \ldots, x_{n_*}$ を選び，

   $$C_i = A_0 A_1^x \cdots A_n^{x^n} = E(f_A(x_i), r_A(x_i))$$

   で与えられる x_i における評価値のコミットメント $C_1, \ldots, C_{n_*}$ と $x_1, \ldots, x_*$ を送る．

2. ユーザは，乱数 s を選び，

   $$D_i = C_i^{sf_B(x_i)} = E(sF(x_i), r_A(x_i) f_B(x_i))$$

   を求める．n 次のランダムな多項式 $r_B(x) = r'_0 + \cdots + r'_n x^n$ を作り，コミットメント

   $$Q_i = E(sf_B(x_i), r_B(x_i)) = g^{sf_B(x_i)} h^{r_B(x_i)}$$

   を求める．サーバに， $D_1, \ldots, D_{n_*}$, Q_1, $\ldots, Q_{n_*}$ を送る．

3. $y_B = sf_B(x_i)$ とおいて，その知識を有していることを連言のゼロ知識証明

   $$PK\{y_B, r_B \mid D_i = C_i^{y_B} \wedge Q_i = g^{y_B} h^{r_B}\}$$

   で示す．

4. サーバは，任意の値 β と $X_1 = \{x_1, \ldots, x_{n_*-1}\}$ と $X_2 = \{x_2, \ldots, x_{n_*}\}$ について，

   $$E_1 = E(sF(\beta)) = \prod_{i=1}^{n_*-1} D_i^{\Lambda(i, X_1, \beta)} \quad (3)$$

   $$E_2 = E(sF(\beta)) = \prod_{i=2}^{n_*} D_i^{\Lambda(i, X_2, \beta)}$$

   を計算し，$E_1 = E_2$ であれば，F が t 個以上の重根を持っていることを確信する．

5. 更に，ゼロ知識証明を検証し，D_i と Q_i の指数が等しいことを確認する．最後に，（重根のない）ラグランジェの補間式を用いて，

   $$\tilde{E_1} = E(f_B(\beta), r_B(\beta)) = \prod_{i=1}^{n+1} Q_i^{L(i, \tilde{X_1}, \beta)}$$

   $$\tilde{E_2} = E(f_B(\beta), r_B(\beta)) = \prod_{i=2}^{n+2} Q_i^{L(i, \tilde{X_2}, \beta)}$$

   により，β におけるコミットメント $E(f_B(\beta))$ を，$\tilde{X_1} = \{x_1, \ldots, x_{n+1}\}$ と $\tilde{X_2} = \{x_2, \ldots, x_{n+2}\}$ から求め，両者が一致すれば，$f_B(x)$ は n 次以上の多項式である．ただし，$L(i, X, \beta)$ はラグランジェ係数である．

式 (4) $E_1 = E(sF(\beta), sr_A(\beta) f_B(\beta))$ は，準同型性により

$$sF(\beta) = \sum_{i=0}^{n} sF(x_i) \Lambda(i, X_1, \beta))$$

$$sr_A(\beta) f_B(\beta) = \sum_{i=0}^{n_*} sr_A(x_i) f_B(x_i) \Lambda(i, X_1, \beta)$$

と展開できる．F も $f_A f_B$ のいずれも，次数 $n_* - 1$ 以下の多項式であり，命題 2.1 により $E_1 = E_2$ を得る．

3.3 提案方式 2

登録フェーズは方式 1 と同じ．認証フェーズを次に示す．

1. 方式 1 に同じ．

2. ユーザは $i = 1, \ldots, n_*$ について，t 以上の重根を持つ多項式を作る多項式 $f_B(x)$ を知っていることをゼロ知識証明で $PK\{f_B(x_i) \mid$

   $$\prod_{i=1}^{n_*-1} C_i^{\Lambda(i, X_1, \beta) f_B(x_i)} = \prod_{i=2}^{n_*} C_i^{\Lambda(i, X_2, \beta) f_B(x_i)}\}$$

   を証明する．ここで，X_1, X_2, Λ は方式 1 と同様．

3. 最後にユーザは，$PK\{f_B(x_i) \mid$

   $$C_i^{\Lambda f_B(x_i)} \wedge Q_i = g^{f_B(x_i)} h^{r_B} \wedge \tilde{E_1} = \tilde{E_2}\}$$

   により自分の多項式が n 次以上であることをゼロ知識証明する．

3.4 数値例

登録時と認証時の生体情報をそれぞれ，$n = 3$ の $A = \{1, 3, 5\}$, $B = \{1, 5, 7\}$ とする．A, B を根として持つ多項式の積 $F(x)$ は，

$$F(x) = (x-1)^2 (x-5)^2 (x-3)(x-7)$$

となり，$t = 2$ 個の重根を持つ $2n = 6$ 次の多項式である．このとき，$n_* = 2n - t + 2 = 6$ 個の元を持つ集合 $X = \{10, 12, \ldots, 14, 15\}$ の任意の $n_* - 1$ 個の部分集合 $X_1 \subset X$ から，$F(\beta)$ は一意に決まる．命題 2.1 により，$F(0) =$

$$
\begin{aligned}
&= \Lambda(10, x_1, 0)F(10) + \cdots + \Lambda(14, x_1, 0)F(14) \\
&= \Lambda(11, x_2, 0)F(11) + \cdots + \Lambda(15, x_2, 0)F(15)
\end{aligned}
$$

ここで，係数は事前計算が可能であり，例えば，

$$
\begin{aligned}
&\Lambda(10, X_1, 0) \\
&= \left(\frac{-11}{10-11}\right)^2 \left(\frac{-12}{10-12}\right)^2 \frac{-13}{10-13} \frac{-14}{10-14}
\end{aligned}
$$

である．

4 評価

4.1 安全性

提案方式は，重根のサイズ t の範囲内で，あいまい性を満たし，ただしい利用者ならば必ず証明を成功させることが出来る（完全性）．不正な利用者が証明を成功させたならば，その不正者は t 個以上の正しい根を同定選んでいるか，検証式を満足する不正な f_B を逆算できるかのどちらかである（健全性）．検証者には，コミットメントの安全性によって，利用者の生体情報は見えないが，方式 1 では任意の点 β についての検査式を自由に計算できるので，認証情報と登録情報の差 t を調べることが可能である．方式 2 ではそれも防止している．

4.2 認証精度

提案方式の認証精度は，基本となる幾何学的ハッシュの性能に依存する．ゼロ知識証明に関する部分は，セキュリティパラメータを操作して任意の精度で偽造が防止が出来るからである．

そこで，実際の指紋を用いて，提案方式の認証精度を実験的に検証する．実験には NFIS2 `mindtct` プログラムを用いて，研究室内で抽出したマニューシャを用いた．マニューシャ抽出品質を指定して，約 50 個に絞ったマニューシャを用い，端点分岐点の区別は省略した．幾何学的ハッシュのコアは，マニューシャが分布する方形の中心から上位 k までを用いている．マニューシャの向きを注意深くスキャンすることで，回転の線形変換は省略した．

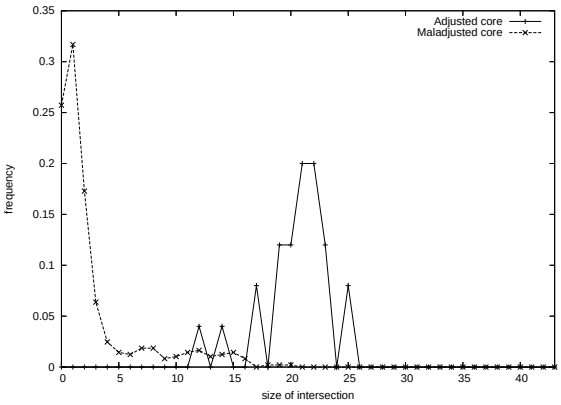

図 1: 幾何学的ハッシュによる共通集合の大きさ

(実験 1) まず，複数回スキャンした同一の指紋を用いて，二つのマニューシャ集合の積集合の大きさを調べた結果を図 1 に示す．$n = 43$ の登録マニューシャ集合 A と 8 個の指紋 B の組について，それぞれ $k = 8$ 回コアを変えて照合を行った．すなわち，$8 \times 8 \times 8 = 512$ 回の組の中から，A と B のコアが一致した時 (adjusted core)，不一致 (maladjusted core) の積集合の大きさの分布を表している．量子化レベルは $q = 20$ で測定した．

図より，サイズ 1 と 20 を中心とした明確に分離できる二つの分布が観測できる．この図より，しきい値 $t = 15$ 位に設定すれば，十分他人と本人の判別が出来そうなことがわかる．(なお，他人の指紋についての分布を調べたところ，この不一致の分布と変わらない平均値であった.)

(実験 2) 必要なコアの試行回数を調べるために，$k = 1, \ldots, 8$ について，認証の成功率を調べた.「認証成功」は，実験 1 より，k 以内のコアのどれかについて，$t \geq 15$ の積集合があること，と定めた．量子化レベルは $q = 20$ である．実験結果を，図 2 に示す.

指紋を単一の登録指紋と照合したときの平均成功率 (average) と，k 個の登録指紋を使ったときの成功率 (total) を示している．k を増やすに従って，比較する組は k^2 生じるので，$k = 4$ で既

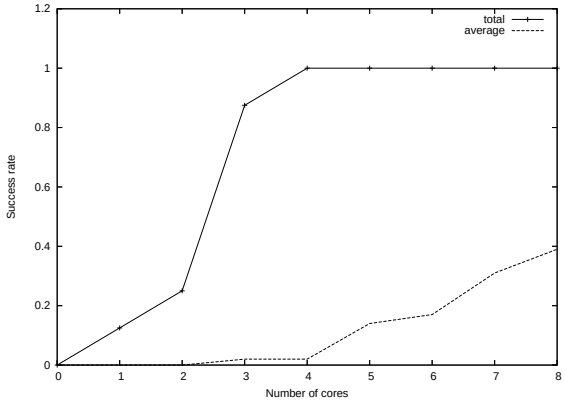

図 2: コアの数に対する認証成功率

表 2: 認証精度

q	20	30	35	40	total
FAR	0.73	0.57	0.42	0.38	100
FRR	0.0	0.0	0.0	0.15	520

に 100%の成功率に達している．従って，$k = 4$ で十分コアの一致が期待できることが分かった．

(実験 3) 12 名からそれぞれ 10 個づつの指紋を採取して，FAR, FRR を求めた．$k = 4, t = 15$ はこれまでの実験とあわせた．量子化レベル q を変動させて，それぞれの誤り率を測定した．結果を，表 2 に示す．

4.3 計算量コスト

表 3 はプロトコルから見積もった提案方式のパフォーマンスである．計算コストは，最も支配的なべき乗計算の数で定めている．

この結果より，二つの提案方式にはパフォーマンスに関する大きな差は生じないことがわかる．また，$n_* = 2n - t + 2 = O(n)$ なので，最も支配的な計算処理は認証ステップ 1 における多項式の評価であり，$2n_*n = 4n^2 - 2tn + 2n$ である．幾何学的ハッシュの実験によると，$n = 40$ 程度であるため，$4 \times 40^2 = 6400$ のべき乗がかかり，一回 10ms で実行すると約 0.6 秒かかる．ただし，この処理は事前計算が可能であり，高速化の可能性を残す．

表 3: 計算量

(方式 1)

step	ユーザ	サーバ
登録 1	$2n$	
認証 1		$2n_*n$
2,4	$3n_*$	$2n_* - 2$
3,5	$3n_*$	$5n_* + 2n$
計	$6n_*$	$2n_*n + 7n_* + 2n - 2$

(方式 2)

step	ユーザ	サーバ
登録 1	$2n$	
認証 1		$2n_*n$
2	n_*	$2n_* - 1$
3	$3n_*$	$5n_* + 2n$
計	$4n_*$	$2n_*n + 7n_* + 2n - 1$

5 結論

持ち物なしで非対称生体認証を行うプロトコルを提案した．生体情報の大きさ n に対して，$O(n^2)$ の計算コストがかかる．

参考文献

[1] E. Fujisaki, T. Okamoto, “Statistical Zero Knowledge Protocols ro Prove Modular Polynomial Relations,” Prpc. of CRYPTO ’99, Springer LNCS vol. 1666, pp. 413–430, (1999)

[2] S. Lee, et al., “Protecting Secret Keys with Fuzzy Fingerprint Vault Based on a 3D Geometric Hash Table,” LNCS 4432, pp. 432-439, 2007.

[3] D. Moon, S. Lee and Y. Chung, “Configurable fuzzy fingerprint vault for Match-on-Card System”, IEICE Electronics Express (ELEX), Vol. 6, No. 14, pp. 993-999, 2009.

[4] A. Juels, M. Sudan, “A Fuzzy Vault Scheme”, International Symposium on Information Theory, pp. 408, *IEEE*, Lausanne, Switzerland, 2002.

[5] U. Uludag, S. Pankanti and A. Jain. “Fuzzy Vault for Fingerprints”, Proc. of Audio- and Video-based Biometric Person Authentication (AVBPA) 2005, pp. 310–319, Rye Brook, NY, July 2005.

[6] 高橋健太，比良田真史，三村昌弘，手塚悟，“セキュアなリモート生体認証プロトコルの提案”, 情報処理学会論文誌，Vol. 49, No. 9, pp. 3016-3027, 2008.