T2R2 東京工業大学リサーチリポジトリ
Tokyo Tech Research Repository

論文 / 著書情報
Article / Book Information

論題	区間のZKIPを用いた生体認証方式の改良
Title	Improvement of the biometric authentication system using ZKIP
著者	尾形わかは, 菊池浩明, 西垣 正勝
Author	Wakaha Ogata, Hiroaki Kikuchi, Masakatsu Nishigaki
掲載誌/書名	第30回情報理論とその応用シンポジウム (SITA2007) 予稿集, Vol. , No. , pp. 689-693
Journal/Book name	The 30th Symposium on Information Theory and its Applications (SITA2007), Vol. , No. , pp. 689-693
発行日 / Issue date	2007, 11

Powered by T2R2 (Tokyo Institute Research Repository)

The 30th Symposium on Information Theory
and its Applications (SITA2007)
Kashikojima, Mie, Japan, Nov. 27–30, 2007

区間の ZKIP を用いた生体認証方式の改良
Improvement of the biometric authentication system using ZKIP

尾形わかは * 菊池 浩明 † 西垣 正勝 ‡
Wakaha Ogata Hiroaki Kikuchi Masakatsu Nishigaki

Abstract— In SITA2006, the authors proposed a biometric authentication system using ZKIP. In this study, we discuss how we make the system more efficient. We show that the system can be improved by storing partial biometric information in smart cards without any sacrifice of security and privacy.

Keywords— biometrics, cryptographic protocol, zero-knowledge interactive proof

1 はじめに

近年，生体情報を用いた方式が個人認証方式として実用化されている．生体認証方式は，パスワード認証における忘却，持ち物認証における紛失がないという点において利便性が高いが，生体情報の取り扱いにおいて課題が多い．第一に，プライバシー情報である生体情報を認証サーバに登録するという行為そのものに抵抗を感じるユーザが多い．第二に，生体情報は更新することができないため，生体情報が流出してしまうと二度とその生体情報を使用できなるという大きな問題をはらんでいる．後者に対しては，近年，生体情報として比較的容易に流出しやすい指紋認証よりも流出しにくい静脈認証などが注目を浴びているが，検証者が生体情報自体を管理している限り，検証者からの流出の可能性がある．

これに対し，Ratha らはキャンセラブルという概念を導入し，画像ブロック置換，マニューシャ非線形変換などの方式を提案した [1]．キャンセラブルな方式では，乱数要素を用いて加工した生体情報をテンプレートとして検証者に登録することによって，生体情報の流出時には登録情報を取り消すことができる．また，サーバに登録されるテンプレートは加工後の生体情報であるため，プライバシーの問題も緩和される．その後，キャンセラブル方式として [2, 3, 4] などが提案されたが，これらの方式では，認証時の通信を盗聴し再送することにより成りすましが可能であり（リプレイアタック），十分な安全性を持つとは言えない．

これに対し筆者らは，区間の零知識証明を利用した安全なキャンセラブル生体情報認証システムを提案している [5]．このシステム（以降，SITA06 方式と呼ぶ）では，通信量・計算量が非常に大きいという欠点を持つ代わりに，成りすましに対して安全で，かつサーバには何の情報も漏らさない．また，高橋，比良田は，テンプレート作成時に用いた乱数を知っていることを示す零知識証明によって，リプレイアタックを防止することを提案している [6]．

本研究では，これまで提案されている方式の中で最高の安全性を持つ SITA06 方式について，IC カード内に保存する情報を増やすことで，通信量・計算量を削減することができることを示す．また，IC カード内に保存する情報からは個人識別の情報が一切漏れず，SITA06 方式と同様のプライバシーや安全性を保障できることを示す．

2 準備

2.1 生体認証方式のモデル

被認証者（ユーザ）と検証者（サーバ）がネットワークを介して認証を行うリモート認証の場合，生体情報のみの認証では安全性を確保しづらいため，IC カードと組み合わせた 2 重の認証が一般的である．生体認証と IC カードによるリモート認証では，一般には IC カードがローカルで生体認証を行い，サーバは IC カードのみを認証する．この場合，IC カード内に生体情報が格納されるため，生体情報の秘匿性が IC カードの耐タンパー性に依存してしまう．本研究では，IC カードの耐タンパー性を仮定せず，サーバが生体情報と IC カードの双方を認証するモデルを研究対象とする．

任意の生体情報はベクトル $\mathbf{x} = (x_1, \ldots, x_m)$ で表わすことができる．生体認証方式は，あらかじめ $\mathbf{x}$ をテンプレートとして登録しておき，認証時には新たに取得した生体情報 $\mathbf{x}' = (x'_1, \ldots, x'_m)$ をテンプレートと比較して差異が小さい場合に限り本人であるとみなす．生体情報の種類によって比較手法・差異の計測手法が異なり，多くの研究者によって誤認証率（FAR，FRR）を小さく抑える手法の開発がされているが，何らかの変数変換を行うことにより，$\mathbf{x}$ と $\mathbf{x}'$ のユークリッド距離が閾値以下である場合に本人であると判定するように一般化できる．このとき，認証の検査式は

$$|\mathbf{x} - \mathbf{x}'|^2 = (x_1 - x'_1)^2 + \cdots + (x_m - x'_m)^2 \leq \theta^2 \quad (1)$$

* 東京工業大学 東京都目黒区大岡山, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo, Japan
† 東海大学 神奈川県平塚市北金目, Tokai University, Kitakaname, Hiratsuka, Kanagawa
‡ 静岡大学 静岡県浜松市城北,Shizuoka University, Johoku, Hamamatsu, Japan,

で表わされる．

この検査式は，いくつかの線形不等式によって近似することができる．例えば，最も単純な近似としては，

$$\forall i : |x_i - x'_i| \leq \theta \tag{2}$$

が挙げられる．さらに良い近似としては，式 (2) に加え

$$|(x_1 - x'_1) \pm (x_2 - x'_2) \pm \cdots \pm (x_m - x'_m)| \leq \sqrt{m}\theta$$

を検証することが考えられる．これを一般化すると，あらかじめ定められたいくつかの係数ベクトル $(a_1^{(j)}, \ldots, a_m^{(j)})$ によって

$$\forall j : |a_1^{(j)}(x_1 - x'_1) + a_2^{(j)}(x_2 - x'_2) + \cdots + a_m^{(j)}(x_m - x'_m)| \leq \theta$$

で書き表すことができる．また，

$$z_j = (a_1^{(j)}, \ldots, a_m^{(j)})\mathbf{x}^T \tag{3}$$

$$z'_j = (a_1^{(j)}, \ldots, a_m^{(j)})(\mathbf{x}')^T \tag{4}$$

と定義すれば，認証の検査式は

$$\forall j : |z_j - z'_j| \leq \theta$$

で書ける．

以降では，登録時に得られたスカラー量 x と，認証時に得られたスカラー量 x' が

$$|x - x'| \leq \theta$$

を満たしていることを検証することを主に考え，後に一般化した近似検査式

$$\forall j : |z_j - z'_j| \leq \theta$$

へ拡張することを考える．

2.2 SITA2006 方式の概要 [5]

キャンセラブル方式では生体情報を乱数によってマスクしテンプレートとするが，SITA06 方式ではマスクとしてコミットメントを使用する．具体的には，整数 x のコミットメントを

$$Com(x, r) = g^x h^r \bmod n$$

とする [7]．ただし，n は誰も素因数を知らない十分大きな合成数，g, h は誰も離散対数 $\log_g h$ を知らないような Z_n^* の要素，r は $[-2^s n + 1, 2^s n - 1]$ から選ばれた乱数，s はセキュリティパラメータ（例えば $s = 40$）である．このコミットメントは準同型性

$$Com(x_1, r_1) \times Com(x_2, r_2) = Com(x_1 + x_2, r_1 + r_2)$$

を満たす．なお，Z_n^* の位数は誰にも分らないので，$x_1 + x_2$ などは mod 演算ではない．

以下に，SITA06 方式の主な流れを示す．

生体情報の登録時には，ユーザは生体情報 x のコミットメント $E = Com(x, r)$ を計算し，これをテンプレートとしてサーバに登録する．また，コミットに用いた乱数 r は IC カード内に保管する．コミットメントの性質より，テンプレートからは生体情報が漏れることはない．

リモート認証の際には，ユーザは取得した生体情報 x' のコミットメント $E' = Com(x', r')$ を計算し，これをサーバに送る．ユーザが本人である場合，2 つのコミットメントにコミットされている 2 つの値は「近い」．そこでユーザは,「コミットされている 2 つの値が十分に近いこと」を証明する零知識証明プロトコルを実行する．具体的には，登録されている生体情報 x の予測値 $\tilde{x} \in \{x', x' \pm 1, \ldots, x' \pm \theta\}$ の全てに対して $\tilde{E} = Com(\tilde{x}, r)$ を計算し，$\tilde{E}$ にコミットされている $\tilde{x}$ と E' にコミットされている x' が十分近い（差が θ 以内である）ことを，区間の零知識証明を用いて証明する．サーバは，$2\theta + 1$ 個の証明のうち，少なくとも 1 つが正しければ認証を受理する．

ここで，区間の零知識証明は，区間に対してログオーダの計算量で実行できるが，$2\theta + 1$ 回の繰り返しを行うため，全体として $O(\theta)$ の計算量が必要となる．

2.3 より良い近似への拡張

生体情報が $\mathbf{x} = (x_1, \ldots, x_m)$ である場合，SITA06 方式を m 回独立に行うことによって

$$\forall i : |x_i - x'_i| \leq \theta$$

を検証することができる．

ここで SITA06 方式を，より良い近似式

$$\forall j : |z_j - z'_j| \leq \theta \tag{5}$$

による検証に拡張する方法について述べる．ここで，z_i, z'_i は $\mathbf{x}$ と $\mathbf{x}'$ から式 (3),(4) によって定義されたものである．

まず，登録時に $\mathbf{x}$ の各要素 x_i からそれぞれコミットメント $E_i = Com(x_i, r_i)$ を計算し，$(E_1, \ldots, E_m)$ をテンプレートとして登録しておく．また，認証時にも $\mathbf{x}'$ の各要素 x'_i からコミットメントを計算しサーバに送信しておく．

コミットメントの準同型性を用いると，E_i から各 z_j のコミットメントを

$$\begin{aligned}\prod_{i=1}^{m} E_i^{a_i^{(j)}} &= \prod_{i=1}^{m} g^{x_i a_i^{(j)}} h^{r_i a_i^{(j)}} \\ &= g^{\sum_{i=1}^{m} x_i a_i^{(j)}} h^{\sum_{i=1}^{m} r_i a_i^{(j)}} \\ &= g^{z_j} h^{t_j}\end{aligned}$$

によって計算できる.（ただし，t_j は乱数 $r_1, \ldots, r_m$ と係数 $a_1^{(j)}, \ldots, a_m^{(j)}$ から計算される値.）同様に E'_i を用いて z'_j のコミットメントも計算できる．したがって，これらのコミットメント E_j, E'_j に基づいて区間の零知識証明を行うことができ，任意の j に対して式 (5) を検証することができる．

2.4 効率化

SITA06 方式では，$2\theta+1$ 回の繰り返しを行うため，非常に効率が悪い．この繰り返しを行う必要があるのは，ユーザが登録した x を全く知らないためである．そこで，x の部分情報のみを IC カードに保管しておくことで効率を改善できることが高橋により指摘されている [8].

以下に，効率化した方式について簡単に説明する．以降では，この方式を SITA06-e 方式と呼ぶ.

登録時には，ユーザは生体情報 x のコミットメント $E = Com(x, r)$ を計算し，これを登録する．また，$y = x \bmod D$ を計算し，(y, r) を IC カード内に保管する．ただし，$D = 2\theta + 1$ である．

認証時には，ユーザはそのときに取得した生体情報 x' のコミットメント $E' = Com(x', r')$ を計算し，これをサーバに送る．また，

$$|\tilde{x} - x'| \leq \theta$$

と

$$y = \tilde{x} \bmod D$$

を共に満たす $\tilde{x}$ を見つける．この $\tilde{x}$ に対して $\tilde{E} = Com(\tilde{x}, r)$ を計算し，$\tilde{E}$ にコミットされている $\tilde{x}$ と E' にコミットされている x' が十分近い（差が θ 以内である）ことを，区間の零知識証明を用いて証明する．サーバは，証明が正しければ正当なユーザと認証する．

D の選び方から，$|\tilde{x} - x'| \leq \theta$ かつ $y = \tilde{x} \bmod D$ を満たす $\tilde{x}$ は一意に定まり，すなわち，それは x そのものである．したがって，y から x を完全に復元できるため，繰り返しを省略することができ，通信量・計算量を $1/(2\theta+1)$ とすることができる．

3 効率化した方式に対する考察

効率化した SITA06-e 方式では，生体情報の部分情報 y を IC カードに保管する．したがって，IC カードの盗難により生体情報が部分的に漏れてしまう．ここでは，この漏れる量を見積もることにより，安全性を考える．

3.1 生体情報の分布に対する仮定

ここでは，個人の生体情報 x は，分散 σ を持つ正規分布に従うと仮定する．分散が大きいほど生体情報の揺らぎが大きいため，認証のためのしきい値 θ と分散 σ は，ある正の係数 c によって $\theta = c\sigma$ の関係にあると考えられる．

FRR を小さくするためには，登録時の生体情報 x と認証時の生体情報 x' の差 $|x - x'|$ が θ 以下である確率を 1 に近くする必要がある．正規分布においては，平均を中心とした 6σ の範囲内に収まる確率が 99.7%程度であることから，c は $c \geq 6$ である必要がある[1].

3.2 漏れる量の見積もり

生体情報の部分情報である y が，どれだけの情報を漏らしているかを測る指標として，相互情報量 $I(U;Y) = H(Y) - H(Y \mid U)$ を用いることにする [9]．ここで，U は全ユーザからランダムにユーザが選ばれる確率変数，Y は，ランダムに取得した生体情報 X から $Y = X \bmod D$ によって計算される確率変数である．$I(U;Y)$ により，ユーザ識別のための生体情報として，何ビットほどの情報漏洩があるのかがわかる．例えば，$I(U;Y) = 0$ は，$H(Y) = H(Y \mid U)$ を意味し，ユーザによらない y の値の散らばり具合と，ユーザを固定したとき y の散らばり具合が等しいことを意味し，したがって y からはユーザの識別は全くできないことを意味する．逆に，$I(U;Y)$ が非常に大きいということは，ユーザを固定したときの y が，固定しないときに比べて極端に偏っている，ということを意味し，y が個人的な情報を含んでいて，これが漏れることは好ましくないことを意味する．

さて，$I(U;Y) = H(Y) = H(Y \mid U)$ を評価する．まず，Y は，ランダムに選ばれたユーザから取得した生体情報 X から $y = x \bmod D$ で計算して得られるものであるから，これは一様分布であるとみなすことができる．したがって，$H(Y) = \log D$ となる．

次に $H(Y \mid U)$ を見積もる．個人を特定したときの生体情報 x は，分散 σ を持つ正規分布に従うと仮定しているので，

$$H(X \mid U) = \log \sqrt{2\pi e \sigma^2}$$

と書ける（e は自然対数の底）．また，$\theta = c\sigma \geq 6\sigma$ かつ $D = 2\theta + 1$ であるので，$D \geq 12\sigma$ であり，このとき，$H(Y \mid U) = H(X \mid U)$ とみなすことができるので，

$$\begin{aligned} I(U;Y) &= \log D - H(X \mid U) \\ &= \log(2\theta+1) - \log\sqrt{2\pi e \sigma^2} \\ &= \log(2\theta+1) - \log\sqrt{2\pi e (\theta/c)^2} \\ &\sim \log \frac{2c}{\sqrt{2\pi e}} \end{aligned}$$

と見積もることができる.（θ によらないことに注意.）

$c = 6$ とした場合，これは 1.54 程度となる．これより，y を IC カードに保存することにより漏洩する情報量は 2 ビ

[1] 登録時に何度か生体情報を取得することによって，x を平均近くにすることができる場合は，$c \geq 3$ となる．

ット未満であることが分かる．これは小さい値ではあるが，FRR を小さく抑えるために c を 6 より大きくした場合にはさらに大きくなることや，生体情報が $\mathbf{x} = (x_1, \ldots, x_m)$ である時には IC カードには $y_i = x_i \bmod D$ で計算される全ての y_i が保存されるため，漏洩する情報も m 倍になることを考えると，SITA06-e 方式は，十分安全な方式であるとは言えない．

4 情報を全く漏らさない効率化の提案

SITA06-e 方式では，x' と y から x を一意に定めるために $D = 2\theta + 1$ とした．しかし，この場合には y から最低 1.54 ビットの情報が漏れることが分かった．

本章では，情報漏洩を完全になくすために，D を小さくすることを考える．

4.1 効率化のアイディア

$I(U;Y)$ を 0 としたいとき，$H(Y) = H(Y \mid U)$ とする必要がある．従って，$H(Y \mid U) = \log D$ となるように D を定めれば良い．一方，X が分散 σ の正規分布に従う場合，$D = 2\sigma$ とすると $Y = X \bmod D$ は一様分布となり，$H(Y \mid U) = \log D$ となる．したがって，$D = 2\sigma = (2/c)\theta$ とすれば，y からの情報漏れを防ぐことができる．

ただし，x' と y から x を一意に定めることはできない．

$$|x - x'| \leq \theta$$

と

$$y = x \bmod D = x \bmod (2/c)\theta$$

を同時に満たす x は c 通りある．したがって，c 個の x の候補全てに対し，区間の ZKIP を実行し，いづれか 1 つが正しければ本人と認証すれば良い．

なお，2.3 節で述べたより良い近似式

$$\forall j : |z_j - z'_j| \leq \theta \tag{6}$$

による検証に拡張した方法についても，全く同様に効率化をすることが可能である．

4.2 効率と安全性の比較

SITA06-e 方式では，IC カードに保管された情報 y から部分情報が漏れるが，提案方式では y からは情報が漏れないため，SITA06 方式と同じ安全性を持つと考えられる．一方，c 個の候補に対して ZKIP を実行する必要があるため，x の候補が 1 つである SITA06-e 方式に比べ，通信量・計算量は c 倍となる．例えば，$c = 6$ の場合には効率は SITA06-e 方式の 1/6 となる．

θ を σ に対して大きくとった場合には候補の数 c は増加するが，θ にはよらないため，SITA06 方式に比べ効率的であると言える．効率と安全性の比較について，表 1 に示す．

表 1: 効率化前の方式との比較

方式	計算量・通信量 *	安全性
SITA06	θ	○
SITA06-e	1	△
本提案	c (≥ 6)	○

* 通信量・計算量は SITA06-e を基準とした値.

ここでは，個人の生体情報は正規分布をすると仮定して解析を行った．したがって，正規分布とは異なる分布をしている場合には，提案方式においても IC カードからの生体情報の一部が漏洩する心配がある．また，正規分布で近似できたとしても，分散 σ を計測することは困難である．そこで，安全性のためには，y の持つ情報量を小さめに設定しておく必要がある．たとえば，生体情報 x の分散の予想値 σ に対して，$D = 2\sigma$ ではなく $D = \sigma$ とすることで，実際の分散が σ より小さい場合にも，情報漏れを防ぐことができる．このとき，$D = (1/c)\theta$ となるため，x の候補は $2c$ 個となり，認証時における通信量・計算量は大きくなるが，効率化前の方式と比較すれば，効率は改善されている．(なお，SITA06 方式は，提案方式において $D = 1$ とした場合の特殊ケースと考えることができる.)

5 まとめと今後の課題

本稿では，[5] において筆者らが提案した方式を基本として，効率化を図った．これにより，漏洩する情報量を 0 に保ちつつ，計算量・通信量を大きく軽減させることができることが分かった．

SITA06 方式や本稿で提案している効率化方式では，線形近似式を使って生体情報が「近い」ことを検証している．しかし，線形近似式の数は $O(2^m)$ であるため，次数 m 大きくなると非常に効率が悪い．とくに，生体情報がバイナリーベクトルであり,「近さ」をハミング距離で表すような生体情報の場合，m は非常に大きな値となるため、実用的とは言えない．これについては，何らかの方策を考える必要がある．

また，他のキャンセラブル方式との効率や安全性の比較を行うことも今後の課題である．

参考文献

[1] N. K. Ratha, K. Karu, S. Chen, and A. K. Jain, “A Real-Time Matching System for Larger Fingerprint Database,” IEEE Trans. on PAMI, Vol.18, No.8, pp.799–813 (1996)

[2] 太田陽基，清本晋作，田中俊昭,「虹彩コードを秘匿する虹彩認証方式の提案」情報処理学会論文誌, Vol.45, No.8, pp.1845–1855 (2004)
[3] 高橋健太，三村昌弘,「キャンセラブル指紋照合方式の提案」コンピュータセキュリティシンポジウム 2005 論文集，pp.379–384 (2005)
[4] 比良田真史，高橋健太，三村昌弘,「画像マッチングに基づく生体認証に適用可能なキャンセラブルバイオメトリクスの提案」2006-CSEC-34, pp.450–440 (2006)
[5] 尾形わかは，菊池浩明，西垣正勝,「リモートバイオメトリクス認証に有効な「近い」ことを示す零知識証明プロトコル」，情報理論とその応用シンポジウム予稿集，SITA2006, pp.319–322 (2006)
[6] 高橋健太，比良田真史,「セキュアなリモート生体認証プロトコルの提案」2007 年暗号と情報セキュリティシンポジウム，SCIS2007 (2007)
[7] E. Fujisaki, T. Okamoto, “Statistical Zero Knowledge Protocols ro Prove Modular Polynomial Relations,” Prpc. of CRYPTO ’99, LNCS vol. 1666, pp. 413–430 (1999)
[8] 高橋，プライベートコミュニケーションによる (2006)
[9] 高橋，日野，村上,「生体情報の情報量に関する一考察」信学技報 Vol.107, No.140, pp.193–200 (2007)