PIRに基づく匿名認証とその応用

中村, 徹
九州大学大学院システム情報科学[府／研究院]

稲永, 俊介
九州大学大学院システム情報科学[府／研究院]

池田, 大輔
九州大学大学院システム情報科学[府／研究院]

馬場, 謙介
九州大学附属図書館研究開発室

他

出版情報：コンピュータセキュリティシンポジウム. 2009, pp.571-576, 2009-10-27
バージョン：
権利関係：

PIR に基づく匿名認証とその応用

中村　徹 †　　稲永　俊介 †　　池田　大輔 †　　馬場　謙介 ‡　　安浦　寛人 †

† 九州大学大学院システム情報科学 [府／研究院]
819-0395 福岡市西区元岡 744 番地

‡ 九州大学附属図書館研究開発室
812-8581 福岡市東区箱崎 6-10-1

{toru,inenaga,yasuura}@c.csce.kyushu-u.ac.jp
daisuke@inf.kyushu-u.ac.jp

baba@lib.kyushu-u.ac.jp

あらまし　本稿では，ユーザ，サービス提供者，データベースの 3 つの主体から構成される認証について注目する．ユーザは認証要求をサービス提供者に送り，サービス提供者は受け取った認証要求を検証する．データベースはサービス提供者に対して，認証要求を検証するために必要な情報を提供する．本稿では，プライバシを考慮した情報獲得手法 (PIR) を用いることにより，(1) データベースに対する匿名性，(2) パスワード保護，(3) 再送攻撃防止，という性質を持つ認証プロトコルを提案する．

Anonymous Authentication Based on PIR and Its Applications

Toru Nakamura† Shunsuke Inenaga† Daisuke Ikeda† Kensuke Baba‡ Hiroto Yasuura†

†Graduate School/Faculty of Information Science and Electrical Engineering, Kyushu University
Moto'oka 744, Nishi-ku, Fukuoka 819-0395, Japan

{toru,inenaga,yasuura}@c.csce.kyushu-u.ac.jp
daisuke@inf.kyushu-u.ac.jp

‡Research and Development Division, Kyushu University Library
10-1, Hakozaki 6, Higashi-ku, Fukuoka, 812-8581, Japan

baba@lib.kyushu-u.ac.jp

Abstract　This paper focuses on authentication with three types of entities: a *user* who sends an authentication request, an *service provider* who receives and verifies the request, and a *database* who supplies the authentication-server with information for verifying the request. This paper presents an authentication protocol which satisfies the following important properties: (1) Anonymity against Database, (2) Password Protection, and (3) Security against Replay Attacks, with a Private Information Retrieval (PIR) scheme.

1　はじめに

認証技術の安全性は，なりすましの脅威を防ぐために重要である．一方で，認証技術のプライバシに関する問題も安全性と同様に高い関心を集めている．実際，ストレージ容量の増大やデータマイニング技術の進歩により，ユーザの通信履歴やサービスログからユーザの行動や嗜好を分析することが容易になっている．

本稿では，3 つの主体から構成される認証について考える．ここでは 3 つの主体を，**ユーザ**，**サービス提供者**，**データベース**と呼ぶ．ユーザは認証要求をサービス提供者に送り，サービス提供者は受け取った認証要求を検証する．デー

タベースはサービス提供者に対して，認証要求を検証するために必要な情報を提供する．本稿では特に，データベースのみが認証に必要な情報を保持し，サービス提供者は全く保持しない場合に注目する．このようなシステムの例として，OpenID [1] など，シングルサインオンシステムがある．このようなシステムの利点として，複数のサービスを利用する場合であっても，一度だけ登録を行えばよい点や，サービス提供者がユーザの情報を漏洩する危険性を削減できる点などがある．しかしながら多くのシステムでは，データベースは，ID や仮名などの情報からサービスを利用しようとするユーザが誰であるか，または同一のユーザであるかどうかを判定することが可能である．ゆえに，データベースは容易にユーザの行動や嗜好を分析することができる．

本稿では，(1) データベースに対してユーザが誰であるか知られることを防ぐ (**データベースに対する匿名性**)，(2) サービス提供者がパスワードを得ることができない (**パスワード保護**)，(3) 再送攻撃を防ぐ (**再送攻撃防止**) という性質を持つ認証プロトコルを提案する．

提案するプロトコルは，プライバシを考慮した情報獲得手法 (以下，PIR) [7, 6, 9, 5] という技術を用いて構成する．PIR とは，データベースに情報獲得の要求を出すユーザのプライバシを保護する技術である．PIR を用いることにより，ユーザはデータベースに対して獲得したい情報のインデックスを秘匿しつつ，その情報を獲得することが可能になる．PIR を用いた認証の既存研究として，PIR を用いた生体認証が提案されている [3, 4]．しかしながら，これらのプロトコルは (1) の性質についてのみ考慮されている．

提案する認証プロトコルは，グループ認証や属性認証を用いるサービスに有効である．例えば，2008 年 7 月から IC カード「taspo」 [2] を用いた成人確認機能付きのたばこの自動販売機が導入された．しかしながら，発行者が個人情報の管理だけでなく，購入者の購入場所や時間の記録を行っているため，個人のプライバシに対する懸念がある．提案するプロトコルを用いることにより，発行者に対して匿名のまま，成人確認を行うことができる．

2 プライバシを考慮した情報獲得手法

データベースに対して獲得したい情報のインデックスを秘匿しつつ，その情報を獲得するナイーブな手法は，ユーザがデータベースに全ての要素を送るように要求することである．このとき，通信量はデータベースの要素数 n に対して $O(n)$ である．Chor ら [7] は，同じ要素を持つ複数のデータベースを用いることにより，情報理論的に安全でかつナイーブな手法と比較して通信量の小さい PIR を提案した．以後，計算量理論的に安全な PIR [6] や，単一のデータベースで実現可能な PIR [9] が提案されている．

PIR 手法の例として，データベース数が 2 の場合 [7] の PIR 手法の概要を説明する．ここでは，説明を簡単にするためにデータベースの要素をビットとし，要素数 n のデータベースをビット列 $X = x_1 \cdots x_n$ とする．(要素がビットではなくビット列である一般的な PIR については，文献 [7] で効率のよい手法が提案されている)．それぞれのデータベースは同じ X を持つ．ユーザはデータベースから i 番目の要素 x_i を獲得したいとする．ユーザは片方のデータベースに $\{1, 2, \ldots, n\}$ の部分集合全体から無作為に選んだ S を送り，もう一方のデータベースに，$i \notin S$ であれば $S \cup \{i\}$ を，そうでなければ $S \setminus \{i\}$ を送る．それぞれのデータベースは，受け取った集合 I に対して，$j \in I$ である全ての x_j の排他的論理和を計算し，それをユーザに送る．これらの 2 つのビットの排他的論理和は，x_i である．データベースが得られる情報はランダムに選択された集合のみであるので，データベースは結託しない限り i に関する何の情報も得ることができない．

本稿では，シングルサーバ PIR を用いる．以下に，[5] に基づくシングルサーバ PIR の定義を示す．ここでは，データベースの要素をビットではなく m ビットのビット列とする．$p(\cdot)$ を任意の多項式とする．任意の自然数 n に対して，

$[n] \stackrel{\text{def}}{=} \{1, 2, ..., n\}$ とする．2 出力のアルゴリズム F が与えられた時，入力 x に対する k 番目の出力を $F_k(x)$ と表す．

定義 1 $m, n, \ell_r, \ell_q, \ell_s, \ell_a \in \mathbf{N}$ について，シングルサーバ *PIR* は以下の *3* つのアルゴリズムから構成される．

- **クエリー生成アルゴリズム** $Q : [n] \times \{0,1\}^{\ell_r} \to \{0,1\}^{\ell_q} \times \{0,1\}^{\ell_s}$
- **アンサー生成アルゴリズム** $A : (\{0,1\}^m)^n \times \{0,1\}^{\ell_q} \to \{0,1\}^{\ell_a}$
- **再構成アルゴリズム** $R : [n] \times \{0,1\}^{\ell_q} \times \{0,1\}^{\ell_s} \times \{0,1\}^{\ell_a} \to \{0,1\}^m$

これらのアルゴリズムは以下の性質を満たす．

- 任意の $X = \{x_i \mid i \in [n], x_i \in \{0,1\}^m\}$，任意の $i \in [n]$ について，

$$\Pr[R(i, Q(i,r), A(X, Q_1(i,r))) = x_i] > 1 - \frac{1}{p(\log n + \ell_q + \ell_s + \ell_a)}$$

  である．ただし，左辺の確率は $\{0,1\}^{\ell_r}$ から一様に選択された r により決まる．

- 任意の $i, j \in [n]$，任意の確率的多項式時間アルゴリズム $\mathcal{B}$，任意の十分に大きい w について，

$$|\Pr[\mathcal{B}(1^w, Q_1(i,r)) = 1] - \Pr[\mathcal{B}(1^w, Q_1(j,r')) = 1]| < \frac{1}{p(w)}$$

  である．ただし，左辺の確率は $\{0,1\}^{\ell_r}$ から一様かつ独立に選択された r, r' 及び $\mathcal{B}$ の動作に用いるランダムな選択 *(*以後，コイントス*)* により決まる．

Q の 1 番目の出力をクエリー，2 番目の出力をシークレット，A の出力をアンサーと呼ぶことにする．

PIR を用いた情報獲得は，具体的には以下のような手順で行う．

1. ユーザは r をランダムに選択し，$(q, s) \leftarrow Q(i, r)$ を計算する．次にデータベースに $q = Q_1(i, r)$ を送る．
2. データベースは $a \leftarrow A(X, q)$ を計算し，ユーザに a を送る．
3. ユーザは $R(i, (q, s), a)$ を計算し，x_i を得る．

3 データベースに対して匿名な認証

本章では，パスワード保護，再送攻撃耐性，データベースに対する匿名性を持つ認証プロトコルを示す．まず安全性要件の定義を行い，次に安全性要件を満たすプロトコルを構成する．プロトコルの構成には，シングルサーバ PIR とチャレンジ･レスポンス認証プロトコルを用いる．

3.1 安全性要件

本論文では以下のような主体から構成される認証モデルを考える．

- **ユーザ:**ユーザ数を n とする．このとき，$i \in [n]$ について，ユーザを $\mathcal{U}_i$ と表す．それぞれのユーザ $\mathcal{U}_i$ には，一意な識別子 i とパスワード $p_i \in \{0,1\}^m$ が割り当てられている．
- **サービス提供者:**サービス提供者 $\mathcal{S}$ は識別子 i についての認証要求を出したユーザが本当にユーザ $\mathcal{U}_i$ であるかどうかを検証する．
- **データベース:**データベース $\mathcal{D}$ は，全てのユーザのパスワードの列 $P = (p_1, p_2, ..., p_n)$ を持つ．P の要素はそれぞれランダムに割り当てられるとする．

プロトコルは，対話チューリング機械を用いてモデル化する．2 つの対話チューリング機械 $\mathcal{A}, \mathcal{B}$ が，互いに対話を行いながら実行する計算をプロトコルと呼び，$\langle \mathcal{A}, \mathcal{B} \rangle$ と表す．$\mathcal{A}, \mathcal{B}$ の入力

として，x, y がそれぞれ与えられた時の出力を $\langle \mathcal{A}(x), \mathcal{B}(y) \rangle$ と表す．3 つの対話チューリング機械からなるプロトコルも 2 つの場合の単純な拡張である．認証プロトコル $\langle \mathcal{P}, \mathcal{V}, \mathcal{M} \rangle$ を構成する確率的多項式時間対話チューリング機械 $\mathcal{P}, \mathcal{V}, \mathcal{M}$ はそれぞれ，ユーザ，サービス提供者，データベースの認証時の動作を表す．$\mathcal{P}$ は入力として識別子 i とパスワード候補 z が与えられ，$\mathcal{M}$ は入力として P が与えられる．$\mathcal{V}$ は実行後に，1/0 を出力する．

以下に，提案プロトコル $\langle \mathcal{P}, \mathcal{V}, \mathcal{M} \rangle$ が満たすべき安全性要件を挙げる．

- **完全性:** 任意の $m, n \in \mathbf{N}$，任意の $i \in [n]$，任意の $P = \{p_i \mid i \in [n], p_i \in \{0,1\}^m\}$ について，

  $$\Pr[\langle \mathcal{P}(i, p_i), \mathcal{V}, \mathcal{M}(P) \rangle = 1] > 1 - \frac{1}{p(mn)}$$

  である．ただし，左辺の確率は $\mathcal{P}, \mathcal{V}, \mathcal{M}$ のコイントスにより決まる．

- **健全性:**任意の $m, n \in \mathbf{N}$，任意の $i \in [n]$，任意の $P = \{p_i \mid i \in [n], p_i \in \{0,1\}^m\}$，任意の $z \neq p_i \in \{0,1\}^m$ について，

  $$\Pr[\langle \mathcal{P}(i, z), \mathcal{V}, \mathcal{M}(P) \rangle = 1] < \frac{1}{p(mn)}$$

  である．ただし，左辺の確率は $\mathcal{P}, \mathcal{V}, \mathcal{M}$ のコイントスにより決まる．

- **パスワード保護:**任意の $m, n \in \mathbf{N}$，任意の $i \in [n]$，任意の確率的多項式時間アルゴリズム $\mathcal{B}$，十分に大きい w について，

  $$\Pr[\mathcal{B}(1^w, t_1) = p_i] < \frac{1}{p(w)}$$

  である．ただし，P の要素は全て独立かつ一様な $\{0,1\}^m$ 上の確率変数であり，t_1 は $\mathcal{P}$ の入力が (i, p_i)，$\mathcal{M}$ の入力が P のときに $\langle \mathcal{P}, \mathcal{V}, \mathcal{M} \rangle$ を実行した場合に $\mathcal{V}$ が得られる情報を表す．左辺の確率は P, t_1 と，$\mathcal{B}$ のコイントスにより決まる．

- **再送攻撃耐性:** $m, n \in \mathbf{N}$，任意の $i \in [n]$，任意の確率的多項式時間アルゴリズム $\mathcal{B}$，十分に大きい w について，

  $$\Pr[\langle \mathcal{B}(1^w, t_2), \mathcal{V}, \mathcal{M}(P) \rangle = 1] < \frac{1}{p(w)}$$

  である．ただし，P の要素は全て独立かつ一様な $\{0,1\}^m$ 上の確率変数であり，t_2 は $\mathcal{P}$ の入力が (i, p_i)，$\mathcal{M}$ の入力が P のときに $\langle \mathcal{P}, \mathcal{V}, \mathcal{M} \rangle$ を実行した場合の $\mathcal{P}, \mathcal{V}$ 間で送受信した情報を表す．左辺の確率は P, t_2 と，$\mathcal{B}$ のコイントスにより決まる．

- **データベースに対する匿名性:**任意の $m, n \in \mathbf{N}$，任意の $i, j \in [n]$，任意の $z, \bar{z} \in \{0,1\}^m$，任意の確率的多項式時間アルゴリズム $\mathcal{B}$，十分に大きい w について，

  $$|\Pr[\mathcal{B}(1^w, t_3) = 1] - \Pr[\mathcal{B}(1^w, t_4) = 1]| < \frac{1}{p(w)}$$

  である．ただし，P の要素は全て独立かつ一様な $\{0,1\}^m$ 上の確率変数であり，t_3, t_4 はそれぞれ，$\mathcal{P}$ の入力が (i, z) で $\mathcal{M}$ の入力が P のときと，$\mathcal{P}$ の入力が $(j, \bar{z})$ で $\mathcal{M}$ の入力が P のときの，$\langle \mathcal{P}, \mathcal{V}, \mathcal{M} \rangle$ を実行した場合に $\mathcal{M}$ が得られる情報を表す確率変数である．左辺の確率は P, t_3, t_4 と，$\mathcal{B}$ のコイントスにより決まる．

3.2 データベースに対して匿名な認証プロトコル

提案するプロトコルは，シングルサーバ PIR (Q, A, R) と，関数族 $H = \{H_k\}_{k \in \mathbf{N}}$(ただし，$H_k : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ とする) を用いたチャレンジ・レスポンスを用いて構成する．ここでは H は以下のような性質を持つと仮定する．

- 任意の $m \in \mathbf{N}$，任意の確率的多項式時間アルゴリズム $\mathcal{B}$，十分に大きい w について，

  $$\Pr[\mathcal{B}(1^w, H_m(x, y)) = x] < \frac{1}{p(w)}$$

かつ，

$$\Pr[\mathcal{B}(1^w, H_m(x,y)) = y] < \frac{1}{p(w)}$$

である．ただし，これらの不等式の左辺の確率は，$\{0,1\}^m$ から一様かつ独立に選択された x, y 及び $\mathcal{B}$ のコイントスにより決まる．

- 任意の $m \in \mathbf{N}$，任意の $x \in \{0,1\}^m$，任意の確率的多項式時間アルゴリズム $\mathcal{B}$，十分に大きい w について，

$$|\Pr[\mathcal{B}(1^w, H_m(x,r)) = 1] - \Pr[\mathcal{B}(1^w, r') = 1]| < \frac{1}{p(w)}$$

である．ただし，左辺の確率は，$\{0,1\}^m$ から一様かつ独立に選択された r, r' 及び $\mathcal{B}$ のコイントスにより決まる．

ここでは，以下のような単純なチャレンジ・レスポンス認証プロトコル $\langle \mathcal{P}, \mathcal{V} \rangle$ を用いる．$\mathcal{P}$ と $\mathcal{V}$ はそれぞれ入力として $x, y \in \{0,1\}^m$ を与えられるとする．

1. $\mathcal{V}$ は $\mathcal{P}$ にランダムチャレンジ $c \in \{0,1\}^m$ を送る．
2. $\mathcal{P}$ はレスポンス $x' \leftarrow H_m(x,c)$ を送る．
3. $\mathcal{V}$ は $y' \leftarrow H_m(y,c)$ を計算し，もし $y' = x'$ であれば 1 を出力し，そうでなければ 0 を出力する．

上記のプロトコルを用いることにより，$\mathcal{P}$ と $\mathcal{V}$ の間の通信路を盗聴してなりすましを行う攻撃を防ぐことができる．

提案プロトコル $\langle \mathcal{P}, \mathcal{V}, \mathcal{M} \rangle$ を以下に示す．

1. $\mathcal{P}$ は $\mathcal{V}$ に $i \in [n]$ を送る．
2. $\mathcal{V}$ は $r \in \{0,1\}^{\ell_r}$ 及び $c \in \{0,1\}^m$ をランダムに選択し，$(q,s) \leftarrow Q(i,r)$ を計算する．$\mathcal{V}$ は $\mathcal{M}$ に (q,c) を送る．
3. $\mathcal{M}$ は，全ての $j \in [n]$ について，$p'_j \leftarrow H_m(p_j, c)$ を計算し，$P' = (p'_1, p'_2, \ldots, p'_n)$ とする．次に $\mathcal{M}$ は $a \leftarrow A(P', q)$ を計算し，$\mathcal{V}$ に a を送る．
4. $\mathcal{V}$ は $p'_i \leftarrow R(i, r, (q,s), a) = H_m(p_i, c)$ を計算する．次に $\mathcal{S}$ は $\mathcal{P}$ に c を送る．
5. $\mathcal{P}$ は $z' \leftarrow H_m(z,c)$ を計算し，$\mathcal{V}$ に z' を送る．
6. もし $z' = p'_i$ であれば，$\mathcal{V}$ は 1 を出力し，そうでなければ 0 を出力する．

上記のプロトコルを図 1 に示す．

定理 1 提案プロトコル $(\mathcal{P}, \mathcal{V}, \mathcal{M})$ は完全性・健全性・パスワード保護・再送攻撃耐性・データベースに対する匿名性を満たす．

証明 証明の概要のみを示す．完全性及び健全性については，定義 1 より明らかである．パスワード保護及び再送攻撃耐性については，背理法を用いて H の性質の仮定との矛盾を導くことができる．データベースに対する匿名性については，背理法を用いて定義 1 との矛盾を導くことができる． □

4 終わりに

本稿では，PIR を用いることにより，パスワード保護，再送攻撃耐性，データベースに対する匿名性を持つ認証プロトコルを提案した．

今後の課題として，本稿での提案プロトコルで実現した性質に加えて，サービス提供者に対する匿名性を実現するプロトコルの構成を考えている．そのアイデアは，公開鍵暗号と，インデックスなしに復元可能な PIR を用いることである．インデックスなしに復元可能な PIR について，調査・提案を行う予定である．

謝辞

本研究の一部は科学技術振興事業団（JST）の戦略的創造研究推進事業（CREST）の支援によるものである．

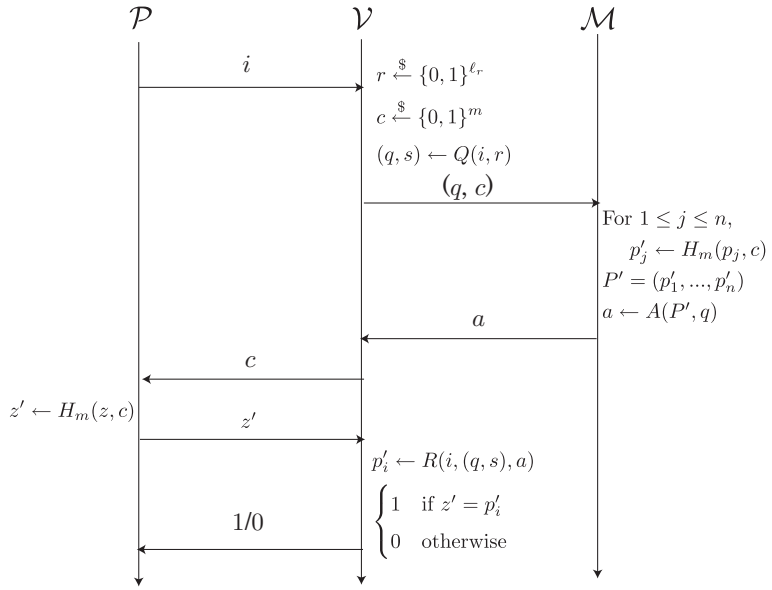

図 1: データベースに対して匿名な認証プロトコル

参考文献

[1] OpenID. http://openid.net/.

[2] Taspo. http://www.taspo.jp/.

[3] Julien Bringer, Hervé Chabanne, Malika Izabachéne, David Pointcheval, Qiang Tang, and Sébastien Zimmer. An application of the Goldwasser-Micali cryptosystem to biometric authentication. In *Information Security and Privacy, 12th Australasian Conference, ACISP 2007*, Vol. 4586 of *LNCS*, pp. 96–106. Springer-Verlag, 2007.

[4] Julien Bringer, Hervé Chabanne, David Pointcheval, and Qiang Tang. Extended private information retrieval and its application in biometrics authentications. In *Cryptology and Network Security, 6th International Conference, CANS 2007*, Vol. 4856 of *LNCS*, pp. 175–193. Springer-Verlag, 2007.

[5] Christian Cachin, Silvio Micali, and Markus Stadler. Computationally private information retrieval with polylogarithmic communication. In *Advances in Cryptology - EUROCRYPT '99*, Vol. 1592 of *LNCS*, pp. 402–414. Springer-Verlag, 1999.

[6] Benny Chor and Niv Gilboa. Computationally private information retrieval. In *Annual ACM Symposium on Theory of Computing*, pp. 304–313. ACM, 1997.

[7] Benny Chor, Oded Goldreich, Eyal Kushilevitz, and Madhu Sudan. Private information retrieval. *Jornal of the ACM*, Vol. 45, pp. 965–982, 1998.

[8] Oded Goldreich. *Foundations of Cryptography*. Cambridge University, 2001.

[9] Eyal Kushilevitz and Rafail Ostrovsky. Replication is not needed: Single database, computationally-private information retrieval. In *the 38th Annual Symposium on Foundations of Computer Science*, pp. 364–373, 1997.