Anonymous Authentication Systems Based on Private Information Retrieval

Nakamura, Toru
Graduate School/Faculty of Information Science and Electrical Engineering, Kyushu University

Inenaga, Shunsuke
Graduate School/Faculty of Information Science and Electrical Engineering, Kyushu University

Ikeda, Daisuke
Graduate School/Faculty of Information Science and Electrical Engineering, Kyushu University

Baba, Kensuke
Graduate School/Faculty of Information Science and Electrical Engineering, Kyushu University

他

出版情報：Networked Digital Technologies. 1, pp.53-58, 2009-07
バージョン：
権利関係：

Anonymous Authentication Systems Based on Private Information Retrieval

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 baba@ait.kyushu-u.ac.jp

Abstract

This paper focuses on authentication with three types of entities: a user *who sends an authentication request, an* authentication-server *who receives and verifies the request, and a* database *who supplies the authentication-server with information for verifying the request. This paper presents novel authentication protocols that satisfy the following important properties: (1) secure against replay attacks, (2) the database(s) cannot identify which user is authenticating and (3) the authentication-server cannot identify to which user a given authentication-request corresponds. Firstly, we show a protocol with a single database which satisfies Properties (1) and (2). Secondly, we show a protocol with multiple databases which satisfies Properties (1), (2) and (3). A key idea of our authentication protocols is to use* private information retrieval *(*PIR*) [Chor et al. J. ACM, 1998].*

1 Introduction

Identity management technologies are getting more and more essential in our life. In particular, authentication plays an important role to prevent impersonation attacks. In the meantime, much attention has been paid to protecting users' privacy on identity management systems [10]. Indeed, due to increase of data storage available and progress of data mining technologies, it is becoming easier for adversaries to analyze user's actions or preferences from e.g., the service logs related to the user.

In this paper, we focus on authentication with three types of entities: a *user* who sends an authentication request, an *authentication-server* who receives and verifies the request, and a *database* who supplies the authentication-server with information for verifying the request. These types of systems are recently becoming more and more popular (e.g., the OpenID system [1] which is a kind of single-sign-on system), since each user needs to register him/her in the system only once even in a multi-service environment, and since the risk of leakage of user-related information from an authentication-server is reduced.

This paper presents novel authentication protocols that satisfy the following important properties: (1) secure against replay-attacks, (2) the database(s) cannot identify which user is authenticating (*anonymity against the database(s)*), and (3) the authentication-server cannot identify to which user a given authentication-request corresponds (*anonymity against the authentication-server*). Firstly, we show a protocol with a single database which satisfies Properties (1) and (2). Secondly, we show a protocol with multiple databases which satisfies Properties (1), (2) and (3).

A key idea of our authentication protocols is to use the *private information retrieval* (*PIR*) [8] technologies. Using the PIR technologies, a client (the authentication-server in our context) can retrieve a data string from a database without the index of the data string being revealed to the database.

Related Work

It should be noted that there already exist authentication protocols based on PIR. Bringer *et al.* [3, 4] proposed biometric authentication protocols based on PIR. However, unfortunately their protocols do not satisfy Properties (1) and (3).

It is known that authentication protocols based on anonymous credentials [6] or group signatures [7] also satisfy Properties (2) and (3). A merit of these protocols is that the authentication-server does not need to communicate with the manager (the database in our context) when verifying a given authentication request. However, a revocation of a group member in these schemes is known to be difficult due to the anonymity. Still, it is easy to revoke a user in our simple password-based authentication protocols.

Liao *et al.* [12] proposed a password-based anonymous authentication protocol. In this protocol, the authentication-

server needs to secretly store some information to verify a given authentication request. On the other hand, in our protocols the authentication-server does not need to store such information.

2 PIR Based Authentication

In this section we propose a simple authentication protocol that is based on private information retrieval (PIR). The simple protocol has properties that

- the authentication-server does not need not to store a set of passwords of users, and
- the database cannot identify which user is authenticating with the authentication-server.

2.1 Security Requirements

We consider an authentication protocol which consists of the three following types of entities:

- **Users**: A *user* $\mathcal{U}_i$ is assigned a unique *identifier* $i \in [n] \stackrel{\text{def}}{=} \{1, 2, ..., n\}$ and has a *password* $p_i \in \{0,1\}^m$.
- **Authentication-Server**: An *authentication-server* $\mathcal{S}$ verifies whether or not an entity who has sent an authentication request with identifier i is truly user U_i.
- **Databases**: A *database* $\mathcal{D}$ stores a set $P = \{p_1, p_2, ..., p_n\}$ of passwords of users.

Throughout this paper, we assume that

- each user can communicate only with the authentication-server,
- the authentication-server can communicate with both users and databases,
- each database can communicate only with the authentication-server,
- the password of each user is shared by the user and a database, and
- the authentication-server have no passwords.

Although this paper only considers protocols with a single authentication-server, it is rather straightforward to apply our protocols to multiple service environments.

In what follows, we regard each user $\mathcal{U}_i$, the authentication-server $\mathcal{S}$, and a database $\mathcal{D}$ as interactive algorithms. A user $\mathcal{U}_i$ takes a pair of an identifier i and a *password candidate* z as input, and a database $\mathcal{D}$ takes P as auxiliary input.

It is important for an authentication protocol to satisfy the following requirements:

- *Correctness*: If $z = p_i$, then the probability that the user $\mathcal{U}_i$ is rejected by the authentication-server $\mathcal{S}$ is negligible.
- *Soundness*: If $z \neq p_i$, then the probability that the user $\mathcal{U}_i$ is accepted by authentication-server $\mathcal{S}$ is negligible.
- *Anonymity against Database*: It is hard for the database $\mathcal{D}$ to compute any information about the identifier i from any information that $\mathcal{D}$ obtains in the authentication protocol.

2.2 Simple Authentication Protocol Based on PIR

In this subsection, we construct a simple authentication protocol based on the single-database PIR [11]. We show a formal definition of a single-database PIR protocol based on the formulation in [5].

Let n be the number of data stored in the database. Let ℓ_r, ℓ_q, ℓ_s, ℓ_a be any non-negative integers. Let $m > 2^n$ be the length of a password p_i. Let $poly(\cdot)$ denote any polynomial.

Definition 1 *A single-database PIR for consists of the following three functions:*

- Query function $Q : [n] \times \{0,1\}^{\ell_r} \to \{0,1\}^{\ell_q} \times \{0,1\}^{\ell_s}$;
- Answer function $A : (\{0,1\}^m)^n \times \{0,1\}^{\ell_q} \to \{0,1\}^{\ell_a}$;
- Reconstruction function $R : [n] \times \{0,1\}^{\ell_q} \times \{0,1\}^{\ell_s} \times \{0,1\}^{\ell_a} \to \{0,1\}^m$.

For integer i and string r of length ℓ_r, let $Q_1(i, \ell_r)$ denotes the first element q of the output $(q, s) = Q(i, r)$. These functions satisfy the following requirements:

- *For any set $X = \{x_i \mid 1 \leq i \leq n, x_i \in \{0,1\}^m\}$ and any $i \in [n]$,*

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

  where the probability is taken over uniformly chosen $r \in \{0,1\}^{\ell_r}$.

- *For any $i, j \in [n]$, any probabilistic polynomial-time algorithm $\mathcal{B}$, and sufficiently large w,*

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

  where the probabilities are taken over uniformly and independently chosen $r, r' \in \{0,1\}^{\ell_r}$ and random coins of $\mathcal{B}$.

Below we show a simple authentication protocol based on PIR (see also Figure 1).

1. $\mathcal{U}_i$ sends a pair of an identifier i and a password candidate z to $\mathcal{S}$ as an authentication request.

2. $\mathcal{S}$ chooses r randomly and computes $(q, s) \leftarrow Q(i, r)$. $\mathcal{S}$ sends q to $\mathcal{D}$.

3. $\mathcal{D}$ computes $a \leftarrow A(P, q)$ and sends a to $\mathcal{S}$.

4. $\mathcal{S}$ computes $p_i \leftarrow R(i, (q, s), a)$ and compares p_i and z. If $z = p_i$, then $\mathcal{S}$ outputs 1, otherwise, $\mathcal{S}$ outputs 0.

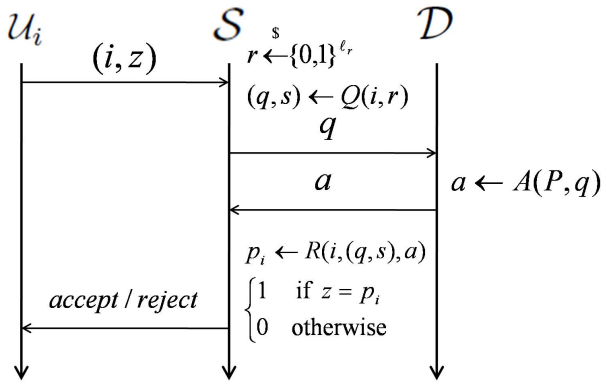

Figure 1. Simple authentication protocol based on PIR.

Theorem 1 *The simple authentication protocol based on PIR satisfies correctness and soundness.*

proof: Due to Inequality (1) of Definition 1, the probability that $\mathcal{S}$ does not receive password p_i from the database $\mathcal{D}$ is negligible. Now assume $\mathcal{S}$ has successfully received p_i. If $z = p_i$, then the probability that $\mathcal{U}_i$ is rejected by $\mathcal{S}$ is negligible. If $z \neq p_i$, then the probability that $\mathcal{U}_i$ is accepted by $\mathcal{S}$ is negligible. Therefore, the simple authentication protocol based on PIR has correctness and soundness. □

Theorem 2 *The simple authentication protocol based on PIR satisfies anonymity against database.*

proof: Note that $q = Q_1(i, r)$ is all the information related to i that $\mathcal{D}$ can obtain in the above protocol. By Definition 1, it is hard for any polynomial-time algorithms to compute any information about i from q. Hence the protocol has anonymity against database. □

We remark that the simple authentication protocol based on PIR has two disadvantages: One is that the password candidate z is transmitted in the communication channel in Step 1. Hence an adversary can impersonate the user $\mathcal{U}_i$ if $z = p_i$ is eavesdropped by the adversary. This cannot be solved by simply encrypting or hashing the password candidate, since an adversary can impersonate the user by repeating the previous communication transcript to the authentication-server (replay attacks). The other is that the authentication-server $\mathcal{S}$ can obtain password p_i in Step 4. This contributes to an increased risk of leakage of user's passwords.

In the following sections, we will show authentication protocols that solve the above difficulties.

3 Authentication Protocol Preventing Replay Attacks

In this section, we propose an authentication protocol based on PIR which prevents the authentication-server from obtaining a password, and prevents replay-attacks. We apply the idea of challenge-response to the simple authentication protocol of the previous section.

3.1 Password Protection and Security against Replay-Attack

In addition to the requirements introduced in the previous section, we consider the two following requirements.

- *Password Protection* : it is hard for the authentication-server $\mathcal{S}$ to compute the user's password p_i from any information that $\mathcal{S}$ obtains in the authentication protocol.
- *Security against Replay-attacks*: it is hard for any adversary who can obtain transcripts of previous communication between $\mathcal{U}_i$ and $\mathcal{S}$ to be accepted by $\mathcal{S}$.

3.2 Challenge-Response Authentication Protocol

Our protocol is based on a challenge-response authentication protocol like CRAM-MD5 [2], which uses a cryptographic hash function. We assume that there exists an ideal hash function such that

- it is hard to guess the input from an output (*one-wayness*),
- it is hard to find two inputs that hash to the same output (*collision resistance*), and
- it is hard to distinguish whether an outputs from the hash function or from true random function (*pseudo-randomness*).

Readers are referred to e.g. [9] for formal definitions of the above properties.

Let $H : \{0,1\}^* \rightarrow \{0,1\}^m$ be an ideal hash function. We denote a concatenation of any strings a and b by $a \parallel b$.

We can construct a *challenge-response authentication protocol* with a single-server PIR and an ideal hash function, as follows (see also Figure 2).

1. $\mathcal{U}_i$ sends an identifier i to $\mathcal{S}$.
2. $\mathcal{S}$ chooses r, r' randomly and independently, and computes $(q, s) \leftarrow Q(i, r)$. $\mathcal{S}$ sends (q, r') to $\mathcal{D}$.
3. For every $1 \leq j \leq n$, $\mathcal{D}$ computes $p'_j \leftarrow H(p_j \parallel r')$, and let $P' = \{p'_1, p'_2, ..., p'_n\}$. $\mathcal{D}$ computes $a \leftarrow A(P', q)$, and sends a to $\mathcal{S}$.
4. $\mathcal{S}$ computes $p'_i \leftarrow R(i, r, (q, s), a) = H(p_i \parallel r')$. $\mathcal{S}$ sends r' to $\mathcal{U}_i$.
5. $\mathcal{U}_i$ computes $z' \leftarrow H(z \parallel r')$, where z is a password candidate, and sends z' to $\mathcal{S}$.
6. If $z' = p'_i$, then $\mathcal{S}$ outputs 1, otherwise, $\mathcal{S}$ outputs 0.

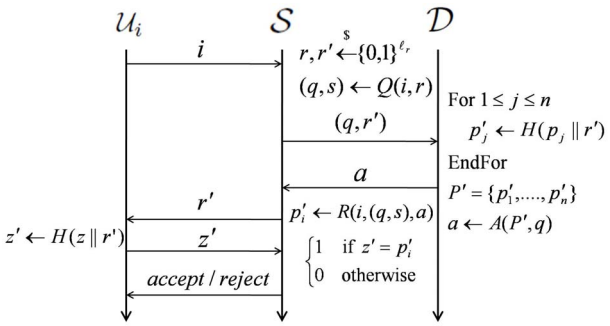

Figure 2. Challenge-Response Authentication Protocol Based on PIR.

Theorem 3 *The challenge-response authentication protocol based on PIR satisfies correctness, soundness, anonymity against database, password protection, and security against replay-attack.*

proof: (Correctness and soundness) Due to Inequality (1) of Definition 1, the probability that $\mathcal{S}$ does not receive password p_i from the database $\mathcal{D}$ is negligible. Now assume $\mathcal{S}$ has successfully received p_i. If $z = p_i$, then clearly $H(z \parallel r') = H(p_i \parallel r')$. Hence, if $z = p_i$, then the probability that $\mathcal{U}_i$ is rejected by $\mathcal{S}$ is negligible, and if $z \neq p_i$, then the probability that $\mathcal{U}_i$ is accepted by $\mathcal{S}$ is negligible.

(Anonymity against Database) In the above protocol, database $\mathcal{D}$ obtains $\{q, r'\}$ such that $q = Q_1(i, r)$. Since r is random value, r clearly includes no information about i. By Definition 1, it is hard for any polynomial-time algorithms to compute any information about i from q. Thus the protocol has anonymity against database.

(Password Protection) The authentication-server $\mathcal{S}$ obtains $\{i, a, p'_i, z'\}$ such that $p'_i = H(p_i \parallel r')$ and $z' = H(z \parallel r')$. The value i is independent of both p_i and z. Since $q = Q_1(i, r)$ and H is a one-way hash function, it is hard to compute p_i from $a = A(P', q)$. Due to the one-wayness of function H, the probability that $\mathcal{S}$ succeeds in computing p_i (or z) from p'_i (or z') is negligible. Therefore the protocol has the property of password protection.

(Security against Replay-attacks) Transcripts of a previous communication between $\mathcal{U}_i$ and $\mathcal{S}$ is $\{i, r', z'\}$. The probability that any adversary succeeds in computing z from z' is negligible, since H has one-wayness and pseudo-randomness. Similarly, it is hard for any adversary to compute $H(z \parallel r'')$ from $H(z \parallel r')$ with a fresh random value r'' with $r'' \neq r'$. Therefore, the protocol has security against replay-attack. □

4 Authentication Protocol Anonymous against Authentication-Server

In this section, we propose an authentication protocol based on PIR in which a user does not have to send his identifier to the authentication-server. Hence can neither the authentication-server or the databases identify which user corresponds to a given authentication-request.

4.1 Anonymity against Authentication-Server

In addition to the four requirements shown in the previous sections, we consider the following requirement.

- *Anonymity against Authentication-Server* : It is hard for the authentication-server $\mathcal{S}$ to compute any information about the identifier i from any information that $\mathcal{S}$ obtains in the authentication protocol.

4.2 PIR with Reconstruction Function Not Taking Identifier as Input

In the single-database PIR of Definition 1, the reconstruction function takes an index of a data string as an input to reconstruct the corresponding data string. In this section, we use an information theoretical PIR protocol of [8] in which the reconstruction function does not take as input an index of a data string. In the information theoretical PIR, there are k database-servers which have the same copies of the database.

Definition 2 *An information theoretical k-database PIR without identifiers in reconstruction consists of the following three functions:*

- k Query function $Q^1, ..., Q^k : [n] \times \{0,1\}^{\ell_r} \to \{0,1\}^{\ell_q}$;
- Answer function $A : (\{0,1\}^m)^n \times \{0,1\}^{\ell_q} \to \{0,1\}^{\ell_a}$;
- Reconstruction function $R : (\{0,1\}^{\ell_a})^k \to \{0,1\}^m$.

These functions satisfy the following requirements:

- *For any set* $X = \{x_i \mid 1 \le i \le n, x_i \in \{0,1\}^m\}$ *and any* $i \in [n]$,

$$R(A(X, Q^1(i,r)), \ldots, A(X, Q^k(i,r))) = x_i.$$

- *For any* $i, j \in [n]$, *any* $t \in [k]$, *and for any* $q \in \{0,1\}^{\ell_q}$

$$\Pr[Q^t(i,r)) = q] = \Pr[Q^t(j,r')) = q],$$

where the probabilities are taken over uniformly and independently chosen $r, r' \in \{0,1\}^{\ell_r}$.

- *For any* $i, j \in [n]$ *and for any* $x \in \{0,1\}^m$,

$$\Pr[R(A(X, Q^1(i,r)), \ldots, A(X, Q^k(i,r))) = x]$$
$$= \Pr[R(A(X, Q^1(j,r)), \ldots, A(X, Q^k(j,r))) = x],$$

where the probabilities are taken over uniformly and independently chosen $r, r' \in \{0,1\}^{\ell_r}$.

4.3 Authentication Protocol Anonymous against Authentication-Server

The key idea of our authentication protocol is to use a *public key encryption scheme*. A public key encryption scheme consists of a triple of algorithms, *key generation algorithm* K, *encryption algorithm* E, and *decryption algorithm* T, satisfies following properties:

- For any $\alpha \in \{0,1\}^*$, $T(sk, E(pk, \alpha)) = \alpha$, where $(pk, sk) \leftarrow K(1^y)$ (y is a security parameter),
- Semantic secure [9] (we omit a formal definition of this property).

In our protocol, we assume that a pair (pk, sk) is precomputed, and all users have obtained pk, and a manager has stored sk secretly.

Our authentication protocol which realizes anonymity against the authentication-server is as follows (see also Figure 3).

1. $\mathcal{U}_i$ chooses r randomly. For every $1 \le d \le k$, $\mathcal{U}_i$ computes $q_d \leftarrow Q^d(i,r)$ and $q'_d \leftarrow E(pk, q_d)$ as an authentication request. $\mathcal{U}_i$ sends $(q'_1, ..., q'_k)$ to $\mathcal{S}$.
2. $\mathcal{S}$ chooses r' randomly. For every $1 \le e \le k$, $\mathcal{S}$ sends (q'_e, r') to $\mathcal{D}_e$.
3. $\mathcal{D}_e$ computes $q_e \leftarrow T(sk, q'_e)$. For every $1 \le j \le n$, $\mathcal{D}_e$ computes $p'_j \leftarrow H(p_j \parallel r')$, and let $P' = \{p'_1, p'_2, ..., p'_n\}$. $\mathcal{D}_e$ computes $a_e \leftarrow A(P', q_e)$ and sends a_e to $\mathcal{S}$.
4. $\mathcal{S}$ computes $p'_i \leftarrow R(a_1, ..., a_n) = H(p_i \parallel r')$. $\mathcal{S}$ sends r' to $\mathcal{U}_i$.
5. $\mathcal{U}_i$ computes $z' \leftarrow H(z \parallel r')$, where z is a password candidate, and sends z' to $\mathcal{S}$.
6. If $z' = p'_i$, then $\mathcal{S}$ outputs 1, otherwise, $\mathcal{S}$ outputs 0.

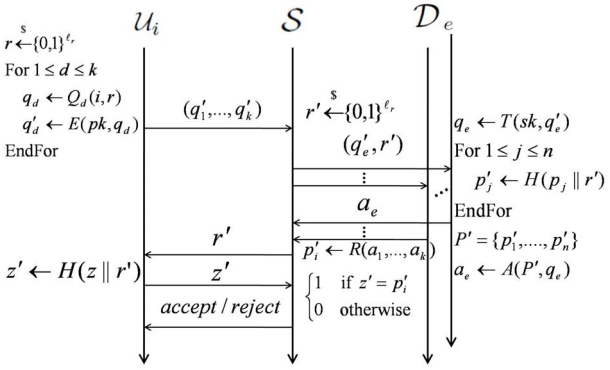

Figure 3. PIR-Based Authentication Protocol Anonymous against Authentication-Server.

Theorem 4 *The proposed protocol satisfies correctness, soundness, password protection, security against replay-attacks, anonymity against databases, and anonymity against authentication-server.*

proof: (Correctness and Soundness) Since we use a public key encryption scheme, every database $\mathcal{D}_e$ can obtain q_e. It follows from Definition 2 that $\mathcal{S}$ can obtain p_i. It is clear that $H(z \parallel r') = H(p_i \parallel r')$ if $z = p_i$. Hence, if $z = p_i$, then $\mathcal{U}_i$ is accepted by $\mathcal{S}$, and if $z \neq p_i$, then $\mathcal{U}_i$ is rejected by $\mathcal{S}$. Therefore, the protocol has correctness and soundness.

(Password Protection and Security against Replay-attacks) These can be shown similarly to Theorem 3 of Section 3.2.

(Anonymity against Database-Servers) In the proposed protocol, each $\mathcal{D}_e$ obtains $\{q_e, r'\}$ such that $q_e = Q^e(i,r)$.

Since r is a random value, r contains no information about i. By Definition 2, it is impossible for any polynomial-time algorithms to compute any information about i from q_e. Thus the protocol has privacy protection.

(Anonymity against Authentication-Server) What $\mathcal{S}$ obtains in the proposed protocol is $\{(q'_1, ..., q'_k), (a_1, ..., a_k), p'_i, z'\}$. $\mathcal{S}$ can obtain no information about i from $(q'_1, ..., q'_k)$ since the public encryption scheme is semantic secure. Also, $\mathcal{S}$ can also obtain no information about i from $(a_1, ..., a_k)$ by Definition 2. The value i is independent from p'_i and z'. Therefore the protocol has anonymity against authentication-servers. □

5 Conclusions and Future Work

In this paper, we proposed novel authentication protocols. Firstly, we showed a protocol with a single database which satisfies correctness, soundness, anonymity against database, password protection, and security against replay-attacks. Secondly, we showed a protocol with multiple databases which satisfies anonymity against authentication-server in addition to the previous properties. The key idea is to employ PIR in the core of authentication protocols.

The authentication protocol proposed in Section 4 is based on an information theoretical PIR in which an identifier is not given as an input of the reconstruction function. However, the assumption that the multiple databases are mutually independent but yet share the same set of passwords, may be impractical. Hence our future work includes surveying or developing single-server PIR without identifiers in reconstruction. We will evaluate the computational complexity and the communication complexity of these proposed protocols.

Acknowledgments

This work was in part supported by CREST-DVLSI of JST. We are grateful for their support.

References

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

[2] RFC2195. http://tools.ietf.org/html/rfc2195.

[3] J. Bringer, H. Chabanne, M. Izabachéne, D. Pointcheval, Q. Tang, and S. Zimmer. An application of the Goldwasser-Micali cryptosystem to biometric authentication. In *Information Security and Privacy, 12th Australasian Conference, ACISP 2007*, volume 4586 of *LNCS*, pages 96–106. Springer-Verlag, 2007.

[4] J. Bringer, H. Chabanne, D. Pointcheval, and Q. Tang. Extended private information retrieval and its application in biometrics authentications. In *Cryptology and Network Security, 6th International Conference, CANS 2007*, volume 4856 of *LNCS*, pages 175–193. Springer-Verlag, 2007.

[5] C. Cachin, S. Micali, and M. Stadler. Computationally private information retrieval with polylogarithmic communication. In *Advances in Cryptology - EUROCRYPT '99*, volume 1592 of *LNCS*, pages 402–414. Springer-Verlag, 1999.

[6] J. Camenisch and A. Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *Advances in Cryptology - EUROCRYPT 2001*, volume 2045 of *LNCS*, pages 93–118. Springer-Verlog, 2001.

[7] D. Chaum and E. van Heyst. Group signatures. In *Advances in Cryptology - EUROCRYPT 1991*, volume 547 of *LNCS*, pages 257–270. Springer-Verlag, 1991.

[8] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan. Private information retrieval. *Jornal of the ACM*, 45:965–982, 1998.

[9] O. Goldreich. *Foundations of Cryptography*. Cambridge University, 2001.

[10] M. Hansen, A. Schwartz, and A. Cooper. Privacy-enhancing identity management. *IEEE Security and Privacy*, 3:38–45, 2008.

[11] E. Kushilevitz and R. Ostrovsky. Replication is not needed: Single database, computationally-private information retrieval. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*, pages 364–373, 1997.

[12] Y.-P. Liao and S.-S. Wang. A secure dynamic ID based remote user authentication scheme for multi-server environment. *Computer standards and interfaces*, 31(1):24–29, 2009.