An Identifiable Yet Unlinkable Authentication System with Smart Cards for Multiple Services

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
Research and Development Division, Kyushu University Library

他

出版情報：Computational Science and Its Applications - Iccsa 2010 : International Conference, Fukuoka, Japan, March 23-26, 2010, Proceedings, Part IV. 6019 (4), pp.236-251, 2010-04-06. Springer
バージョン：

An Identifiable Yet Unlinkable Authentication System with Smart Cards for Multiple Services

Toru Nakamura[1], Shunsuke Inenaga[1], Daisuke Ikeda[1],
Kensuke Baba[2], and Hiroto Yasuura[1]

[1] 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
[2] Research and Development Division, Kyushu University Library
10-1, Hakozaki 6, Higashi-ku, Fukuoka, 812-8581, Japan
baba@lib.kyushu-u.ac.jp

Abstract. The purpose of this paper is to realize an authentication system which satisfies four requirements for security, privacy protection, and usability, that is, *impersonation resistance against insiders*, *personalization*, *unlinkability in multi-service environment*, and *memory efficiency*. The proposed system is the first system which satisfies all the properties. In the proposed system, transactions of a user within a single service can be linked (personalization), while transactions of a user among distinct services can not be linked (unlinkability in multi-service environment). The proposed system can be used with smart cards since the amount of memory required by the system does not depend on the number of services. First, this paper formalizes the property of unlinkability in multi-service environment, which has not been formalized in the literatures. Next, this paper extends an identification scheme with a pseudorandom function in order to realize an authentication system which satisfies all the requirements. This extension can be done with any identification scheme and any pseudorandom function. Finally, this paper shows an implementation with the Schnorr identification scheme and a collision-free hash function as an example of the proposed systems.

1 Introduction

With the increase of the number of services which a user would like to use, it is becoming more and more tedious for the user to establish and manage pairs of a user name (pseudonym) and a password of multiple services. Hence much attention is recently paid to authentication systems which enable users to use multiple services after they register at a registration manager only once. For example, single-sign-on systems, such as Shibboleth [2], OpenID [1], and so on, have been popular. In this paper, such a system is called *an authentication system in multi-service environment*.

The multi-service environment raises a new problem on privacy of users, that is, the daily activity of a user can be revealed from information in multiple service providers. Service providers usually maintain service logs of the transactions for the purpose of

the detection of abuse, audit, and diagnosis of problems, and they can collect their log files and trace actions of a user from his/her transactions. This can be done if the same pseudonym is associated with the same user and is used for multiple service providers. In fact, a typical single-sign-on system is based on such an implementation, hence much more information in various service providers can be collected due to leakages of the service logs or illegal coalitions among multiple service providers. In order to solve the problem, authentication systems should have the property that it is difficult to determine whether multiple transactions in distinct service providers are related to the same user or not (*unlinkability in multi-service environment*). There are some authentication systems which satisfy unlinkability in multi-service environment, such as Janus [6], anonymous credentials [4], and authentication systems based on group signatures [5].

Authentication systems which satisfy unlinkability in multi-service environment can be classified according to the degree of unlinkability as follows.

- Transactions of a user can be linked within a service, while transactions of a user among distinct services. can not be linked.
- Transactions of a user can not be linked even within a service.

From the viewpoint of privacy protection, the systems with the latter property are superior to those with the former property. However, on the practical side, the systems with the latter property have some disadvantages. Indeed, without identification of each user, the purpose of service logs previously described cannot be achieved. Therefore, the system with the latter property cannot be applied to "personalized services", which customize and provide the contents according to a user's profile and preference. Examples of personalized services are personalized news and recommendation services. In the systems with the former property, service providers can identify each user (*Personalizasion*), hence they can maintain the service logs of their users and personalized service can be treated. In this paper, we focus on the systems with the former property.

Next, we consider how to maintain pairs of a pseudonym and a password. There are two ways on how to maintain pairs, that is, (1) doing by himself and (2) delegating the maintenance of the pairs to a trusted third party, such as a registration manager. We focus on the case (1) in this paper. In the case (1), some trusted devices, such as PCs and smart cards, are usually used for storing the pairs. A straightforward solution that satisfies both unlinkability in multi-service environment and personalization is that each user stores the table of the pairs of a pseudonym and a password for all service providers. In this solution, the amount of memory required by the system is proportional to the number of service providers. This solution would be efficient for systems with PCs as they have enough amount of memory. However, in this paper we are interested in situations where the portability of device of a user is indispensable, such as the use of ATM machines. Hence we consider a smart card as a device of a user. Notice that, since smart cards have much less memory than PCs, the above straightforward solution is unsuitable for smart cards when the number of service providers is considerably large. Therefore, it is important for any authentication system with smart cards to require as little amount of memory as possible in order to store pseudonyms and passwords (*Memory efficiency*)

The requirements for an authentication system considered in this paper are the following:

– *Personalization*: service providers can identify each user.
– *Unlinkability in multi-service environment*: it is difficult to determine whether two transactions among distinct service providers are the same user's or not.
– *Memory efficiency*: the amount of memory for pseudonyms and passwords does not depend on the number of service providers.
– *Impersonation resistance against insiders*: even if an adversary is a service provider, the adversary cannot impersonate a legitimate user.

In practical systems, the entities who try to impersonate a legitimate user are not only eavesdroppers but also malicious service providers. Therefore, authentication systems should have the property that an adversary cannot impersonate a legitimate user even if the adversary is a service provider.

We propose the first authentication system which satisfies all the requirements previously described. We note that there is no authentication system satisfies all of the requirements as far as we know. We show an extension of an identification scheme [7], which includes a key generating algorithm and an identification protocol, and the purpose of the extension is to realize the authentication system. The overview of our extended identification protocol is as follows:

– First, a user generates a pair of a pseudonym and a secret-key for each service provider from the corresponding *service ID* with pseudorandom functions [8].
– Next, the user and the service provider follow an identification protocol.

In order to evaluate our extended identification scheme, we define the above requirements based on the computational theory and we prove that our extended identification scheme satisfies all the requirements. To our knowledge, the definition of unlinkability in multi-service environment has not been formalized based on the computational theory, hence we show the first formalization of unlinkability in multi-service environment. The definition of impersonation resistance in this paper is based on the formalization of security of identification schemes in [7].

Related Work

Gabber *et al.* [6] proposed an authentication system, named Janus. In the Janus system, a user generates a pair of a pseudonym and a password for each service provider from his/her secret and the corresponding service ID with a cryptographic function. Hence the amount of memory does not depend on the number of service providers. The Janus system satisfies personarization, unlinkability in multi-service environment, and memory efficiency. However, the property of impersonation resistance was not much treated in [6]. Both Juang [10] and Hwang & Shiau [9] proposed authentication systems in multi-service environment with smart cards which satisfy memory efficiency. However, these systems cannot achieve unlinkability in multi-service environment. Liao and Wang [11] proposed the anonymous authentication system in multi-service environment with smart cards which have both memory efficiency and unlinkability in multi-service environment. However, service providers cannot identify each user in the system. Similarly, in anonymous credential systems [4] and in the systems based on group signatures [5], service providers cannot identify each user.

Organization

This paper is organized as follows. In Section 2 we recall the definition of identification schemes [7] and introduce its slight modification. In Section 3 we consider an extension of identification schemes to the case where there are multiple service providers. We also formalize the property of unlinkability in multi-service environment. Section 4 shows our proposed identification scheme which satisfies impersonation resistance, unlinkability, memory efficiency, and personalization. Section 5 describes an example of implementation of our authentication system based on the Schnorr identification scheme [12]. Section 6 concludes the paper.

2 Identification Scheme

In this paper, we show an extension of an identification scheme which realizes an authentication system which satisfies all the requirements, that is, impersonation resistance against insiders, personalization, unlinkability in multi-service environment, and memory efficiency. In this section, we first show the definition of identification schemes [7]. Next, we discuss the extension of the definition of identification schemes based on the equality of the outputs of protocols.

2.1 Definitions

An *interactive Turing machine* (ITM) is a multi-tape Turing machine with read-only input tapes, a read-and-write work tape, a write-only output tape, a pair of communication tapes, and a read-and-write switch tape consisting of a single cell. One communication tape is read-only and the other is write-only.

Two ITMs $\mathcal{A}$ and $\mathcal{B}$ are said to be linked if

- an input tape of $\mathcal{A}$ coincides with an input of $\mathcal{B}$,
- the read-only communication tape of $\mathcal{A}$ coincides with the write-only communication tape of $\mathcal{B}$, and vice versa, and
- the switch tape of $\mathcal{A}$ coincides with that of $\mathcal{B}$.

The shared input tape is called the *common input tape* of the two ITMs, while the other tapes are called an *auxiliary input tape*. A *joint computation* of two linked ITMs is a sequence of pairs of the local configurations (that is, the state, the contents of the tapes, and the positions of the heads) of the ITMs, where the configuration of one ITM is not modified when the configuration of the other ITM is modified, which is realized by the switch tape (if the content of the switch tape is 0, the configuration of the one ITM is modified, and otherwise that of the another one is modified). The output of a joint computation is the content of the output tape of one of the ITMs.

The output of a Turing machine $\mathcal{A}$ on an input x is denoted by $\mathcal{A}(x)$. We denote by $\langle \mathcal{A}, \mathcal{B} \rangle$ a joint computation of ITMs $\mathcal{A}$ and $\mathcal{B}$, and by $\langle \mathcal{A}(y), \mathcal{B}(z) \rangle (x)$ its output on a common input x, an auxiliary input y for $\mathcal{A}$, and an auxiliary input z for $\mathcal{B}$. We sometimes omit the brackets if the input tapes are blank. In the rest of this paper, we sometimes call a Turing machine $\mathcal{A}$ an "algorithm" $\mathcal{A}$, and a joint computation $\langle \mathcal{A}, \mathcal{B} \rangle$ a "protocol". If $\mathcal{A}$ is a probabilistic algorithm, $\mathcal{A}_r(x)$ denotes the output of $\mathcal{A}$ on an input x and random coins r. We denote by $p(n)$ denotes any polynomial of $n \in \mathbf{N}$.

Definition 1. *An* identification scheme *is a pair of a probabilistic polynomial-time algorithm $\mathcal{I}$ and a protocol $\langle \mathcal{P}, \mathcal{V} \rangle$ of two probabilistic polynomial-time ITMs such that:*

– Viability: *For any $n \in \mathbf{N}$, any $\alpha \in \{0,1\}^n$, and any $s \in \{0,1\}^n$,*

$$\Pr[\langle \mathcal{P}(s), \mathcal{V} \rangle(\alpha, \mathcal{I}_s(\alpha)) = 1] = 1.$$

– Impersonation resistance against insiders: *For any pair $(\mathcal{B}', \mathcal{B}'')$ of probabilistic polynomial-time ITMs, any sufficiently large $n \in \mathbf{N}$, any $\alpha \in \{0,1\}^n$, and any z,*

$$\Pr[\langle \mathcal{B}''(z, T), \mathcal{V} \rangle(\alpha, \mathcal{I}_S(\alpha)) = 1] < \frac{1}{p(n)},$$

where S is a random variable uniformly distributed over $\{0,1\}^n$ and T is a random variable describing the output of $\mathcal{B}'(z)$ after interacting with $\mathcal{P}(S)$, on common input $(\alpha, \mathcal{I}_S(\alpha))$, for polynomially many times.

Then, the string s is called a *secret-key*, the string α is called a *pseudonym*, the algorithm $\mathcal{I}$ is called a *verifying-key generating algorithm*, the output of $\mathcal{I}$ is called a *verifying-key*, and the protocol $\langle \mathcal{P}, \mathcal{V} \rangle$ is called an *identification protocol*.

2.2 Extension Based on Equality of Output of Protocols

In this section, we extend the identification scheme by the equality of the outputs of protocols. We also show the extended identification schemes which satisfies viability and impersonation resistance.

For any protocol $\langle \mathcal{A}, \mathcal{B} \rangle$ and any input x, it is easy to see that there exists a protocol $\langle \mathcal{A}', \mathcal{B}' \rangle$ such that

$$\langle \mathcal{A}, \mathcal{B} \rangle(x) = \langle \mathcal{A}'(x), \mathcal{B}'(x) \rangle.$$

In addition, it is easy to see that there exists a protocol $\langle \mathcal{A}'', \mathcal{B}'' \rangle$ such that

$$\langle \mathcal{A}'(x), \mathcal{B}'(x) \rangle = \langle \mathcal{A}''(x), \mathcal{B}'' \rangle.$$

The next lemma follows from the above arguments.

Lemma 1. *For any identification protocol $\langle \mathcal{P}, \mathcal{V} \rangle$, any $n \in \mathbf{N}$, any $\alpha \in \{0,1\}^n$, and any $s \in \{0,1\}^n$, there exists a protocol $\langle \mathcal{P}', \mathcal{V}' \rangle$ such that*

$$\langle \mathcal{P}(s), \mathcal{V} \rangle(\alpha, \mathcal{I}_s(\alpha)) = \langle \mathcal{P}'(s, \alpha), \mathcal{V}' \rangle(\mathcal{I}_s(\alpha)).$$

For instance, the protocol $\langle \mathcal{P}', \mathcal{V}' \rangle$ can be constructed as follows:

1. $\mathcal{P}'$ is an ITM which reads α on the auxiliary input tape, writes α in the write-only communication tape, and then behaves in the same manner as $\mathcal{P}$.
2. $\mathcal{V}'$ is a modification of $\mathcal{V}$, which reads α on the read-only communication tape instead of reading α on the common input tape.

The modified version of identification protocol $\langle \mathcal{P}', \mathcal{V}' \rangle$ is called the *extended identification protocol* w.r.t. $\langle \mathcal{P}, \mathcal{V} \rangle$.

Lemma 2. *If $(\mathcal{I}, \langle \mathcal{P}, \mathcal{V} \rangle)$ is an identification scheme and $\langle \mathcal{P}', \mathcal{V}' \rangle$ is the extended identification protocol w.r.t. $\langle \mathcal{P}, \mathcal{V} \rangle$, the* extended identification scheme *$(\mathcal{I}, \langle \mathcal{P}', \mathcal{V}' \rangle)$ satisfies the following property: for any pair $(\mathcal{B}', \mathcal{B}'')$ of probabilistic polynomial-time ITMs, any sufficiently large $n \in \mathbf{N}$, any $\alpha \in \{0,1\}^n$, and any z,*

$$\Pr[\langle \mathcal{B}''(z, T, \alpha), \mathcal{V}' \rangle(\mathcal{I}_S(\alpha)) = 1] < \frac{1}{p(n)},$$

where S is a random variable uniformly distributed over $\{0,1\}^n$ and T is a random variable describing the output of $\mathcal{B}'(z)$ after interacting with $\mathcal{P}'(S, \alpha)$, on common input $(\mathcal{I}_S(\alpha))$, for polynomially times.

Proof. Assuming that there exists a pair $(\mathcal{C}', \mathcal{C}'')$ of probabilistic polynomial-time ITMs such that for some $\alpha' \in \{0,1\}^n$, some z', and some polynomial $q(n)$ of n,

$$\Pr[\langle \mathcal{C}''(z', T', \alpha), \mathcal{V}' \rangle(\mathcal{I}_{S'}(\alpha')) = 1] \geq \frac{1}{q(n)},$$

where S' is a random variable uniformly distributed over $\{0,1\}^n$ and T' is a random variable describing the output of $\mathcal{C}'(z')$ after interacting with $\mathcal{P}'(S', \alpha')$ on the common input $(\mathcal{I}_{S'}(\alpha'))$ for polynomially times. We can construct a pair $(\mathcal{D}', \mathcal{D}'')$ of probabilistic polynomial-time ITMs such that:

1. $\mathcal{D}'$ is a modification of $\mathcal{C}'$, which reads α on the read-only communication tape instead of reading α.
2. $\mathcal{D}''$ is an ITM which skips writing α on the write-only communication tape, and then behaves in the same manner as $\mathcal{C}''$.

Then the distribution of the random variable T'', which describes the output of $\mathcal{D}'$ after interacting with $\mathcal{P}(S')$, equals the distribution of T'. According to previous 1 and 2, the pair of $\mathcal{D}'$ and $\mathcal{D}''$ satisfies the following property:

$$\Pr[\langle \mathcal{D}''(z', T'), \mathcal{V} \rangle(\alpha, \mathcal{I}_{S'}(\alpha')) = 1] \geq \frac{1}{q(n)}.$$

This is contradictory to Definition 1. □

The next theorem follows from Lemma 1 and Lemma 2:

Theorem 1. *If $(\mathcal{I}, \langle \mathcal{P}, \mathcal{V} \rangle)$ is an identification scheme and $\langle \mathcal{P}', \mathcal{V}' \rangle$ is the extended identification protocol w.r.t. $\langle \mathcal{P}, \mathcal{V} \rangle$, the extended identification scheme $(\mathcal{I}, \langle \mathcal{P}', \mathcal{V}' \rangle)$ satisfies the following property:*

– Viability*: for any $n \in \mathbf{N}$, any $\alpha \in \{0,1\}^n$, and any $s \in \{0,1\}^n$,*

$$\Pr[\langle \mathcal{P}'(s, \alpha), \mathcal{V}' \rangle(\mathcal{I}_s(\alpha)) = 1] = 1.$$

– Impersonation resistance against insiders*: for any pair $(\mathcal{B}', \mathcal{B}'')$ of probabilistic polynomial-time ITMs, any sufficiently large $n \in \mathbf{N}$, any $\alpha \in \{0,1\}^n$, and any z,*

$$\Pr[\langle \mathcal{B}''(z, T), \mathcal{V}' \rangle(\mathcal{I}_S(\alpha)) = 1] < \frac{1}{p(n)},$$

where S is a random variable uniformly distributed over $\{0,1\}^n$ and T is a random variable describing the output of $\mathcal{B}'(z)$ after interacting with $\mathcal{P}'(S, \alpha)$ on common input $\mathcal{I}_S(\alpha)$, for polynomially many times.

3 Extension of Identification Scheme for Multi-Service Environment and Unlinkability in Multi-Service Environment

In this section, we define identification schemes in multi-service environment by extending identification schemes of Definition 1. The key is the use of a set of functions that map strings to strings. We also formalize the property of *unlinkability in multi-service environment*.

3.1 Extension of Identification Scheme for Multi-Service Environment

In order to describe identification schemes in multi-service environment, we introduce *user IDs* and *service IDs*, which are n-bit strings corresponding uniquely to users and service providers, respectively. We consider a set of functions that map strings which indicate service IDs to strings which indicate pseudonyms or secret-keys. For ease of explanation, we consider only length-preserving functions. Let F be a set of functions that map n-bit strings to n-bit strings, that is, $F = \{f_x : \{0,1\}^n \to \{0,1\}^n\}_{x \in \{0,1\}^n}$, where $x \in \{0,1\}^n$ indicates a user ID.

Let F and G be sets of functions mapping n-bit strings to n-bit strings. For any user ID a and any service ID b, $f_a(b)$ and $g_a(b)$ denote the secret-key and the pseudonym corresponding to the pair (a, b), respectively.

Then, we define *identification schemes in multi-service environment*, which is a quadruplet of a verifying-key generating algorithm $\mathcal{I}$, an identification protocol $\langle \mathcal{P}, \mathcal{V} \rangle$, sets F, and G of functions. An identification scheme in multi-service environment is constructed by replacing a secret-key s and a pseudonym α in Definition 1 with $f_a(b)$ and $g_a(b)$, respectively. An identification scheme in multi-service environment clearly satisfies the property of viability in Definition 1.

3.2 Unlinkability in Multi-Service Environment

We define the property concerning privacy protection, which is called *unlinkability in multi-service environment*. Informally, this property means that it is difficult for any adversaries to determine whether two pseudonyms (and secret-keys) for distinct service IDs are generated from the same user ID or not. We define this property as follows:

Definition 2. *An identification scheme in multi-service environment* $(\mathcal{I}, \langle \mathcal{P}, \mathcal{V} \rangle, F, G)$ *has* unlinkability in multi-service environment *if for any probabilistic polynomial-time algorithm* $\mathcal{A}$*, any sufficiently large* $n \in \mathbf{N}$*, and any* $b \neq b' \in \{0,1\}^n$*,*

$$\Pr[\mathcal{A}(g_U(b), g_U(b')) = 1] - \Pr[\mathcal{A}(g_U(b), g_W(b'))] = 1| < \frac{1}{p(n)}$$

and

$$|\Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_{f_U(b')}(g_U(b'))) = 1] - \Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_{f_W(b')}(g_W(b'))) = 1]| < \frac{1}{p(n)},$$

where U *and* W *are random variables independently and uniformly distributed over* $\{0,1\}^n$*.*

As an example of "linkable" schemes, we consider an identification scheme in multi-service environment in which the same secret-key and pseudonym (we assume they are unique for each user ID) are used for all the service providers. That is, we assume that for any a, f_a and g_a are functions which output the same string on any input b, that is, for any $a \in \{0,1\}^n$ and any $b, b' \in \{0,1\}^n$, $f_a(b) = f_a(b')$ and $g_a(b) = g_a(b')$. In this scheme, it is trivial to check whether or not two pseudonyms for distinct service providers are related to the same user. If an algorithm $\mathcal{A}'$ outputs 1 if the first input equals the second input, and outputs 0 otherwise, it then holds that $\Pr[\mathcal{A}'(g_U(b), g_U(b')) = 1] = 1$ and $\Pr[\mathcal{A}'(g_U(b), g_W(b')) = 1] < 1/p(n)$. Hence this scheme does not have unlinkability in multi-service environment.

4 Identification Scheme Achieving Impersonation Resistance, Unlinkability, Memory Efficiency, and Personalization

In this section, we propose an identification scheme in multi-service environment which satisfies impersonation resistance against insiders, unlinkability in multi-service environment, memory efficiency on auxiliary input tape, and personalization by using an identification scheme and pseudorandom functions [8].

4.1 Proposed Scheme

We explain the overview of our proposed scheme. Assume that each user stores two functions to generate his/her pseudonyms and secret-key. First, after receiving a service ID, a user generates the pair of the pseudonym and the secret-key with the service ID and his/her functions. Next, the user and the corresponding service provider follow an identification protocol. In order to evaluate our scheme, we further modify the definition of identification schemes.

Extension of Identification Scheme for Construction of Our Scheme For any function f, let $\langle f \rangle$ be the description of an algorithm which on an input x returns $f(x)$, and we assume any Turing machine can execute the algorithm which compute f if the machine is given the description $\langle f \rangle$. If $\langle \mathcal{P}, \mathcal{V} \rangle$ is an identification protocol and $\langle \mathcal{P}', \mathcal{V}' \rangle$ is the extended identification protocol w.r.t. $\langle \mathcal{P}, \mathcal{V} \rangle$, *the re-extended identification protocol* $\langle \mathcal{P}'', \mathcal{V}' \rangle$ w.r.t. $\langle \mathcal{P}, \mathcal{V} \rangle$ is constructed as follows:

- $\mathcal{P}''$ is an ITM which first reads $\langle f_a \rangle$ and $\langle g_a \rangle$ on the auxiliary input tape. After reading b on the common input tape, $\mathcal{P}''$ computes $f_a(b)$ and $g_a(b)$. Next, $\mathcal{P}''$ reads $f_a(b)$ and $g_a(b)$ instead of reading the auxiliary input s, α of $\mathcal{P}'$, and then behaves in the same manner as $\mathcal{P}'$.

Our proposed scheme is a quadruplet of a verifying-key generating algorithm $\mathcal{I}$, a re-extended identification protocol $\langle \mathcal{P}'', \mathcal{V}' \rangle$, pseudorandom functions F, and G. In what follows, we show that our identification scheme satisfies impersonation resistance against insiders, unlinkability in multi-service environment, memory efficiency on auxiliary input tape, and personalization.

Pseudorandom Functions A *pseudorandom function*, which is a multi-set of functions that map strings to strings, cannot be distinguished from a truly random function.

An *oracle machine* is a Turing machine with an additional tape, called the oracle tape, and two special states, called oracle invocation and oracle appeared. For configurations with states different from oracle invocation, the next configuration is defined as usual. Let γ be a configuration in which the state is oracle invocation, the oracle is a function f, and the contents of the oracle tape is q. Then the configuration following γ is identical to γ, except that the state is oracle appeared, and the content of the oracle tape is $f(q)$. For any oracle machine $\mathcal{M}$ and function f, let $\mathcal{M}^f$ denote the output of $\mathcal{M}$ when given access to the oracle f. The string q is called $\mathcal{M}$*'s query* and $f(q)$ is called the *oracle reply*.

Definition 3. *A multi-set* $F = \{f_x : \{0,1\}^n \to \{0,1\}^n\}_{x \in \{0,1\}^n}$ *is called a* pseudorandom function*, if for any probabilistic polynomial-time oracle machine* M*, and any sufficiently large* $n \in \mathbf{N}$*,*

$$|\Pr[\mathcal{M}^{f_U}(1^n) = 1] - \Pr[\mathcal{M}^H(1^n) = 1]| < \frac{1}{p(n)},$$

where U *is a random variable uniformly distributed over* $\{0,1\}^n$ *and* H *is a random variable uniformly distributed over all functions from* $\{0,1\}^n$ *to* $\{0,1\}^n$*.*

The following three lemmas are used to prove the impersonation resistance and unlinkability in multi-service environment of our identification scheme. The next lemma follows from Definition 3.

Lemma 3. *For any pseudorandom functions* F*, any* $b \in \{0,1\}^n$*, any probabilistic polynomial-time algorithm* $\mathcal{A}$*, and any* $x \in \{0,1\}^n$*,*

$$|\Pr[\mathcal{A}(f_U(b), x) = 1] - \Pr[\mathcal{A}(W, x) = 1]| < \frac{1}{p(n)}$$

where U *and* W *are random variables independently and uniformly distributed over* $\{0,1\}^n$*.*

Proof. According to Definition 3, the oracle reply which is given by the random variable H distributed over all functions on any query is obviously a random variable uniformly distributed over $\{0,1\}^n$. Assuming for contrary that there exists a probabilistic polynomial-time algorithm $\mathcal{A}'$ such that for some $x' \in \{0,1\}^n$ and some polynomial $q(n)$ of n,

$$|\Pr[\mathcal{A}'(f_U(b'), x') = 1] - \Pr[\mathcal{A}'(W, x') = 1]| \geq \frac{1}{q(n)}.$$

Let $\mathcal{M}'$ be a probabilistic polynomial-time oracle machine which receives the oracle reply $f_U(b')$ on a query b' and then invokes $\mathcal{A}'$ on inputs $f_U(b')$ and x'. Then we have that

$$|\Pr[\mathcal{M}'^{f_U}(1^n) = 1] - \Pr[\mathcal{M}'^{H}(1^n) = 1]| \geq \frac{1}{q(n)},$$

which contradicts Definition 3 of pseudorandom functions. □

The following lemma can be shown similarly to Lemma 3.

Lemma 4. *For any pseudorandom function F, any $b \in \{0,1\}^n$, any probabilistic polynomial-time algorithm $\mathcal{A}, \mathcal{B}$, and any $x \in \{0,1\}^n$,*

$$|\Pr[\mathcal{A}(\mathcal{B}(f_U(b), x)) = 1] - \Pr[\mathcal{A}(\mathcal{B}(W, x)) = 1]| < \frac{1}{p(n)},$$

where U and W are random variables independently and uniformly distributed over $\{0,1\}^n$.

The following lemma can be shown similarly to Lemma 4 since any joint computation can be simulated by a probabilistic polynomial-time algorithm.

Lemma 5. *For any pseudorandom functions F and G, any $b \in \{0,1\}^n$, any probabilistic polynomial-time algorithm $\mathcal{A}$, any protocol $\langle \mathcal{B}, \mathcal{C} \rangle$ of probabilistic polynomial-time ITMs, and any $x \in \{0,1\}^n$,*

$$|\Pr[\mathcal{A}(\langle \mathcal{B}(f_U(b), x), \mathcal{C} \rangle (g_U(b), y)) = 1] - \Pr[\mathcal{A}(\langle \mathcal{B}(W, x), \mathcal{C} \rangle (X, y)) = 1]| < \frac{1}{p(n)},$$

where U, W, and X are random variables independently and uniformly distributed over $\{0,1\}^n$.

4.2 Evaluation of Impersonation Resistance

In this section, we show that our proposed scheme in multi-service environment using pseudorandom functions as F and G satisfies impersonation resistance against insiders.

First, we prove that an identification scheme in multi-service environment with pseudorandom functions has impersonation resistance against insiders.

Theorem 2. *For any identification scheme in multi-service environment $(\mathcal{I}, \langle \mathcal{P}, \mathcal{V} \rangle, F, G)$ such that F and G are pseudorandom functions, any pair $(\mathcal{B}', \mathcal{B}'')$ of probabilistic polynomial-time ITMs, any sufficiently large $n \in \mathbf{N}$, any $b \in \{0,1\}^n$, and any z,*

$$\Pr[\langle \mathcal{B}''(z, T'), \mathcal{V} \rangle (g_U(b), \mathcal{I}_{f_U(b)}(g_U(b))) = 1] < \frac{1}{p(n)},$$

where U is a random variable uniformly distributed over $\{0,1\}^n$ and T' is a random variable describing the output of $\mathcal{B}'(z)$ after interacting with $\mathcal{P}(f_U(b))$, on common input $(g_U(b), \mathcal{I}_{f_U(b)}(g_U(b)))$, for polynomially many times.

Proof. For any probabilistic algorithm $\mathcal{A}$, there exists a deterministic algorithm $\mathcal{A}'$ that outputs $\mathcal{A}'(r, x) = \mathcal{A}_r(x)$ on input x and random coins r. According to Lemma 4 and Lemma 5, it holds that for any probabilistic polynomial-time algorithm $\mathcal{A}$ and any $b \in \{0,1\}^n$,

$$|\Pr[\mathcal{A}(\langle \mathcal{P}(f_U(b)), \mathcal{B}' \rangle (g_U(b), \mathcal{I}_{f_U(b)}(g_U(b)))) = 1] - \Pr[\mathcal{A}(\langle \mathcal{P}(W), \mathcal{B}' \rangle (g_U(b), \mathcal{I}_W(g_U(b)))) = 1]| < \frac{1}{p(n)},$$

where U and W are random variables uniformly and independently distributed over $\{0,1\}^n$. Therefore, it holds that

$$|\Pr[\langle\mathcal{B}''(z,T'),\mathcal{V}\rangle(g_U(b),\mathcal{I}_{f_U(b)}(g_U(b)))=1] - \Pr[\langle\mathcal{B}''(z,T),V\rangle(g_U(b),\mathcal{I}_W(g_U(b)))=1]| < \frac{1}{p(n)},$$

where $T = \langle\mathcal{P}(W),\mathcal{B}'\rangle(g_U(b),\mathcal{I}_W(g_U(b)))$ and $T' = \langle\mathcal{P}(f_U(b)),\mathcal{B}'\rangle(g_U(b),\mathcal{I}_{f_U(b)}(g_U(b)))$. According to the definition of impersonation resistance against insiders in Definition 1,

$$\Pr[\langle\mathcal{B}''(z,T),V\rangle(g_U(b),\mathcal{I}_W(g_U(b)))=1] < \frac{1}{p(n)},$$

hence

$$\Pr[\langle\mathcal{B}''(z,T'),V\rangle(g_U(b),\mathcal{I}_{f_U(b)}(g_U(b)))=1] < \frac{1}{p(n)}.$$

□

In the case where the common tape includes b in addition, it is can be proven that the scheme has impersonation resistance against insiders. The next theorem follows from Theorem 1 and Theorem 2:

Theorem 3. *If a pair $(\mathcal{I},\langle\mathcal{P},\mathcal{V}\rangle)$ is an identification scheme, then our proposed identification scheme, which is a quadruplet of $\mathcal{I}$, re-extended identification protocol $\langle\mathcal{P}'',\mathcal{V}'\rangle$ w.r.t. $\langle\mathcal{P},\mathcal{V}\rangle$ and pseudorandom functions F, and G, satisfies the following properties:*

– Viability*: for any $n\in\mathbf{N}$, any $a\in\{0,1\}^n$ and any $b\in\{0,1\}^n$,*

$$\Pr[\langle\mathcal{P}'(\langle f_a\rangle,\langle g_a\rangle),\mathcal{V}'\rangle(b,\mathcal{I}_{f_a(b)}(g_a(b)))=1]=1.$$

– Impersonation resistance against insiders*: for any pair $(\mathcal{B}',\mathcal{B}'')$ of probabilistic polynomial-time ITMs, any sufficiently large $n\in\mathbf{N}$, any $b\in\{0,1\}^n$ and any z,*

$$\Pr[\langle\mathcal{B}''(z,T'),\mathcal{V}'\rangle(b,\mathcal{I}_{f_U(b)}(g_U(b)))=1] < \frac{1}{p(n)},$$

where U is a random variable uniformly distributed over $\{0,1\}^n$ and T' is a random variable describing the output of $\mathcal{B}'(z)$ after interacting with $\mathcal{P}'(\langle f_U\rangle,\langle g_U\rangle)$, on common input b and $\mathcal{I}_{f_U(b)}(g_U(b))$, for polynomially many times.

4.3 Evaluation of Unlinkability in Multi-Service Environment

In this section, we prove that our proposed identification schemes satisfies unlinkability in multi-service environment.

Theorem 4. *Our proposed identification scheme $(\mathcal{I},\langle\mathcal{P}'',\mathcal{V}'\rangle,F,G)$ has unlinkability in multi-service environment.*

Proof. According to Lemma 3,

$$|\Pr[\mathcal{A}(g_U(b), g_U(b')) = 1] - \Pr[\mathcal{A}(g_U(b), X) = 1]| < \frac{1}{p(n)} \tag{1}$$

and

$$|\Pr[\mathcal{A}(g_U(b), g_W(b')) = 1] - \Pr[\mathcal{A}(g_U(b), Y) = 1]| < \frac{1}{p(n)}, \tag{2}$$

where U, W, X, and Y are random variables independently and uniformly distributed over $\{0,1\}^n$. X and Y follow the same distribution, hence

$$|\Pr[\mathcal{A}(g_U(b), X) = 1] - \Pr[\mathcal{A}(g_U(b), Y) = 1]| < \frac{1}{p(n)}. \tag{3}$$

According to Inequalities (1), (2), and (3),

$$|\Pr[\mathcal{A}(g_U(b), g_U(b')) = 1] - \Pr[\mathcal{A}(g_U(b), g_W(b'))] = 1| < \frac{1}{p(n)}. \tag{4}$$

In a similar way, according to Lemma 4,

$$|\Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_{f_U(b')}(g_U(b'))) = 1] - \Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_X(Y)) = 1]| < \frac{1}{p(n)} \tag{5}$$

and

$$|\Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_{f_W(b')}(g_W(b'))) = 1] - \Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_Z(Q)) = 1]| < \frac{1}{p(n)}, \tag{6}$$

where U, W, X, Y, Z and Q are random variables independently and uniformly distributed over $\{0,1\}^n$. X and Z follow the same distribution and Y and Q follow the same distribution, hence

$$|\Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_X(Y)) = 1] - \Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_Z(Q)) = 1]| < \frac{1}{p(n)}. \tag{7}$$

According to Inequalities (5), (6), and (7),

$$|\Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_{f_U(b')}(g_U(b'))) = 1] - \Pr[\mathcal{A}(\mathcal{I}_{f_U(b)}(g_U(b)), \mathcal{I}_{f_W(b')}(g_W(b'))) = 1]| < \frac{1}{p(n)}.$$

□

4.4 Memory Efficiency on Auxiliary Input Tape

The auxiliary input tape of $\mathcal{P}''$ of our proposed identification scheme corresponds to the memory of each smart card of our authentication system. The memory efficiency on auxiliary input tape of identification schemes is defined as follows:

Definition 4. *An identification scheme in multi-service environment* $(\mathcal{I}, \langle \mathcal{P}, \mathcal{V} \rangle, F, G)$ *has* memory-efficiency on auxiliary input tape *if the length of the auxiliary input tape of* $\mathcal{P}$ *is independent of the number of service providers.*

There exist algorithms such that they compute f_a and g_a and the length of their descriptions is independent of the number of service providers. Therefore, we have the following theorem:

Theorem 5. *Our proposed identification scheme* $(\mathcal{I}, \langle \mathcal{P}'', \mathcal{V}' \rangle, F, G)$ *has memory-efficiency on auxiliary input tape.*

4.5 Personalization

The property of personalization is defined as follows:

Definition 5. *An identification scheme in multi-service environment* $(\mathcal{I}, \langle \mathcal{P}, \mathcal{V} \rangle, F, G)$ *has* personalization*, if for any sufficiently large* $n \in \mathbf{N}$ *and any* $b \in \{0,1\}^n$*,*

$$\Pr[f_U(b) = f_W(b)] < \frac{1}{p(n)} \text{ and } \Pr[g_U(b) = g_W(b)] < \frac{1}{p(n)}$$

where U *and* W *are uniformly and independently distributed over* $\{0,1\}^n$*.*

If F and G are pseudorandom functions, the scheme clearly has personalization because of the property of pseudorandom functions.

Theorem 6. *Our proposed identification scheme* $(\mathcal{I}, \langle \mathcal{P}'', \mathcal{V}' \rangle, F, G)$ *has personalization.*

5 An Example of Implementation

In this section, we show an example of implementation of our authentication system. The implementation is based on the Schnorr identification scheme [12], and uses a collision-free hash function instead of pseudorandom functions. We then estimate an overhead with respect to the run time of our scheme.

5.1 The Schnorr Identification Scheme

As an example of identification schemes, we introduce the scheme proposed by Schnorr [12]. The scheme is a three-move identification scheme based on the discrete logarithm problem. Bellare and Paracio [3] showed that the scheme is secure on the assumption that

the one more inversion problem for discrete logarithm is hard in terms of an interactive computation.

The verifying-key generating algorithm in the Schnorr identification scheme outputs (p, q, g, X) on input $s \in \{0,1\}^k$ for a security parameter $k \in \mathbf{N}$, where p is a prime number such that $2^{k-1} \leq p < 2^k$, q is a prime divisor of $p-1$, g is a generator of a subgroup of $\mathbf{Z}_p^*$ of order q, and X is $g^s \bmod p$. In our system, (p, q, g) is regarded as common parameters and (p, q, g) can be computed independently of X. Hence the verifying-key generating algorithm can be divided into the algorithm $\mathcal{C}$, which outputs (p, q, g) on input 1^k, and the algorithm $\mathcal{I}'$, which outputs X on input s.

The Schnorr identification protocol is shown as follows and in Fig.1.

1. $\mathcal{P}$ chooses $y \in \mathbf{Z}_q$ randomly, computes $g^y \bmod p$, and send the result as Y to $\mathcal{V}$;
2. $\mathcal{V}$ chooses $c \in \mathbf{Z}_q$ randomly and sends c to $\mathcal{P}$;
3. $\mathcal{P}$ computes $y + cs \bmod q$ and sends the result as z to $\mathcal{V}$;
4. $\mathcal{V}$ outputs 1 if $g^z = YX^c \pmod p$, and 0 otherwise.

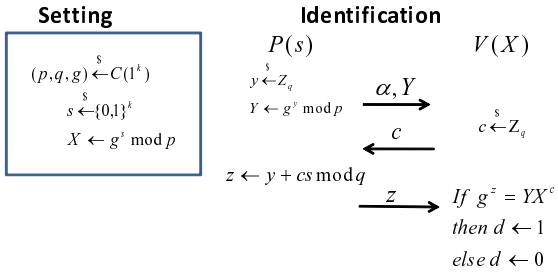

Fig. 1. The Schnorr identification scheme.

5.2 The Implementation of Our Authentication System

Let $\{u_1, u_2, \ldots, u_\ell\}$ be the set of *users* and $\{s_1, s_2, \ldots, s_m\}$ the set of *service providers*. Each user secretly stores his/her user ID in his/her smart card. Let $\{a_1, a_2, \ldots, a_\ell\}$ be the set of user IDs. A user u_i is associated with his/her user ID a_i, and if $i \neq j$, then $a_i \neq a_j$. Each service provider is labelled by his/her service ID, which is the public identifier. Let $\{b_1, b_2, \ldots, b_m\}$ be the set of the service IDs. A service provider s_j is associated with his service ID b_j, and if $i \neq j$, then $b_i \neq b_j$.

We use a collision-free hash function in place of pseudorandom functions. More concretely, $h(0 \parallel a \parallel b)$ and $h(1 \parallel a \parallel b)$ are used as $f_a(b)$ and $g_a(b)$ in the system respectively, where h denotes a collision-free hash function and $\parallel$ denotes concatenation.

In our authentication system, there is a *manager* M, which sets up several parameters. First, we show the preparation procedure which is operated by M.

– **Startup:** M chooses a security parameter $k \in \mathbf{N}$, and computes (p, q, g) with the algorithm $\mathcal{C}$.
– **Registration of Users:** When a new user u_i requests to join in the system, M issues a smart card which stores $a_i \in \{0,1\}^k - \{a_1, a_2, \ldots a_{i-1}\}$ chosen randomly and (p, q, g) to u_i.
– **Registration of Services:** When a new service provider s_j requests to join in the system, M sends $b_j \in \{0,1\}^k - \{b_1, b_2, \ldots b_{i-1}\}$ chosen randomly and (p, q, g) to s_j. Then M computes pairs $(h(0 \parallel a_i \parallel b_j), h(1 \parallel a_i \parallel b_j))$ for all i, and sends pairs $(h(1 \parallel a_i \parallel b_j), g^{h(0\|a_i\|b_j)} \bmod p)$ for all i to s_j.

Next, we show the identification protocol as follows and in Fig. 2.

1. u_i sends an authentication query to s_j.
2. s_j sends b_j to u_i.
3. u_i computes a pair $(h(0 \parallel a_i \parallel b_j), h(1 \parallel a_i \parallel b_j))$ and sends $h(1 \parallel a_i \parallel b_j)$ to s_j.
4. s_j specifies the corresponding $g^{h(0\|a_i\|b_j)} \bmod p$ from $h(1 \parallel a_i \parallel b_j)$.
5. u_i and s_j follow the Schnorr identification scheme.

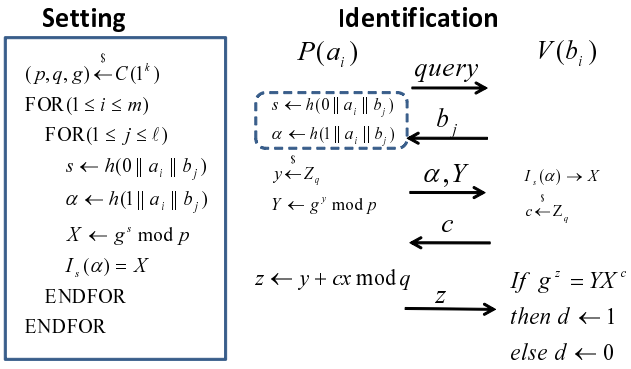

Fig. 2. Our identification scheme based on the Schnorr identification scheme.

5.3 Discussion

A naive scheme realizing unlinkability in multi-service environment can be achieve by storing a user's secret-keys and pseudonyms, which are randomly chosen, in a table. In the naive scheme, the amount of memory which a user needs is proportional to the number of services. Using our scheme, the amount of memory does not depend on the number of services, however, two more hash computations are required compared to the naive scheme. Let $t_a(b) = h(0 \parallel a \parallel b)$ and $t'_a(b) = h(1 \parallel a \parallel b)$. Assuming that the sets $\{t_a\}_{a\in\{0,1\}^n}$ and $\{t'_a\}_{a\in\{0,1\}^n}$ are pseudorandom functions, the implementation of our authentication system satisfies impersonation resistance against insiders, unlinkability in multi-service environment, memory efficiency, and personalization.

6 Conclusions

In this paper we proposed an authentication system in multi-service environment which satisfies impersonation resistance, unlinkability in multi-service environment, memory efficiency, and personalization. Due to the use of pseudorandom functions, the memory requirement for each smart card is independent of the number of services. This is a remarkable advantage when a massive number of services utilize the system. We showed an example of our system based on the Schnorr identification scheme, in which pseudorandom functions are replaced with collision-free hash functions.

Our future work includes the following:

- Implementing our system with smart cards and a PC in order to measure the execution time and to compare it with that of other related authentication systems.
- A comparison of the circuit size of our implementation with a hash function and that of a naive method using a table of pairs of a pseudonym and a secret-key. We conjecture that as the number of service providers increase, our system will become more memory efficient than the naive method.

Acknowledgements

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

References

1. OpenID. http://openid.net/
2. Shibboleth. http://shibboleth.internet2.edu/
3. Bellare, M., Palacio, A.: GQ and schnorr identification schemes: Proofs of security against impersonation under active and concurrent attacks. In: Advances in Cryptology - CRYPTO 2002. LNCS, Springer-Verlog (2002)
4. Camenisch, J., Lysyanskaya, A.: An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In: Advances in Cryptology - EUROCRYPT 2001. LNCS, vol. 2045, pp. 93–118. Springer-Verlag (2001)
5. Chaum, D., van Heyst, E.: Group signatures. In: Advances in Cryptology - EUROCRYPT 1991. LNCS, vol. 547, pp. 257–270. Springer-Verlag (1991)
6. Gabber, E., Information, C., Gibbons, P.B., Matias, Y., Mayer, A.: How to make personalized web browsing simple, secure, and anonymous. In: Financial Cryptography. LNCS, vol. 1318, pp. 17–31. Springer-Verlag (1997)
7. Goldreich, O.: Foundations of Cryptography. Cambridge University (2001)
8. Goldreich, O., Goldwasser, S., Micali, S.: How to construct random functions. Journal of the ACM (JACM) 33(4), 792–807 (1986)
9. Hwang, R.J., Shiau, S.H.: Provably efficient authenticated key agreement protocol for multi-servers. The Computer Journal 50, 602–615 (2007)
10. Juang, W.S.: Efficient multi-server password authenticated key agreement using smart cards. IEEE Transactions on Consumer Electronics 50, 251–255 (2004)
11. Liao, Y.P., Wang, S.S.: A secure dynamic ID based remote user authentication scheme for multi-server environment. Computer standards and interfaces 31(1), 24–29 (2009)
12. Schnorr, C.P.: Efficient signature generation by smart cards. In: Journal of Cryptology. vol. 4, pp. 161–174. Springer New York (1991)