An Identifiable Yet Unlinkable Authentication System in Multi-Service Environment

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

Yasuura, Hiroto
Graduate School/Faculty of Information Science and Electrical Engineering, Kyushu University

他

出版情報：情報処理学会研究報告. MPS, 数理モデル化と問題解決研究報告. 2009 (10), pp.1-7, 2010-03-05. The Information Processing Society of Japan (IPSJ)
バージョン：

An Identifiable Yet Unlinkable Authentication System in Multi-Service Environment

TORU NAKAMURA ,[†1] SHUNSUKE INENAGA ,[†2]
DAISUKE IKEDA ,[†2] KENSUKE BABA [†3]
and HIROTO YASUURA[†2]

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*, *weak-unlinkability*, 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 (weak-unlinkability). 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 weak-unlinkability, 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.

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 only once at a registration manager. For example, credit card systems and single-sign-on systems such as OpenID[10], 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 by gathering information about his/her usage of 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. Service logs of a user in a service provider can be associated with those in other service providers if a pseudonym of the user is used among multiple service providers. In fact, credit card systems and OpenID are typically based on such an implementation ⋆1, 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. For example, it is technically viable that purchase information of credit cards in a supermarket is shared with other companies or is sold to another company without asking the customer. OpenID enables a third party to gather information about user behaviour over multiple web sites without user consent more easily than using "Web Bugs"[1]. Web Bugs are invisible third party images added to a web page so that the third party receives notice of the page viewing event.

In order to solve the problem, authentication systems should have the property that it is difficult to determine whether two transactions in distinct service providers are related to the same user or not (*weak-unlinkability*). There are some authentication systems which satisfy weak-unlinkability, such as Janus[4], anonymous credentials[2], and authentication systems based on group signatures[3].

Authentication systems which satisfy weak-unlinkability 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.

†1 Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan
†2 Faculty of Information Science and Electrical Engineering, Kyushu University, Japan
†3 Research and Development Division, Kyushu University Library, Japan
⋆1 The real author is the Editorial Board of the Trans. on SIGMPS, IPSJ.

⋆1 As a example of the exception, OpenID.ee[12] provides distinct pseudonyms of a user for each web site.

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.

A appropriate strategy to realize weak-unlinkability depends on the way how a user maintains pairs pseudonym and a password. We consider the way that a user maintains the pairs. 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 weak-unlinkability 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 an efficient solution to realize weak-unlinkability 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. 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. For example, VISA, which is a company that offers credit card services, provides the assurance of acceptance at more than 30 million Merchant outlets in 2009[14]. In another instance, the number of web sites which support OpenID is about 50 thousand in July 2009[11]. 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*).

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 (impersonation resistance against insiders).

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

- *Personalization*: service providers can identify each user.
- *Weak-unlinkability*: 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.

We propose an extension of an *identification scheme*[5] to realize the first authentication system which satisfies all the requirements previously described. Identification schemes consist of a key-generating algorithm and an interactive identification protocol. Identification schemes are required that a prover is always able to convince a verifier who the prover is, whereas any adversary cannot impersonate the prover even after getting interactions between the prover and the verifier. We note that there is no authentication system satisfies all of the requirements as far as we know. 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 *service ID* corresponding to the service which the user wants to use by pseudorandom functions[6].
- Next, the user and the service provider follow an identification protocol.

In order to evaluate our identification scheme, we define some properties, which correspond to the above requirements concerning authentication systems, based on the computational theory. We also prove that our identification scheme satisfies all the properties. To our knowledge, the definition of weak-unlinkability has not been formalized based on the computational theory, hence we introduce the first formalization of weak-unlinkability. The definition of impersonation resistance in this paper is based on the formalization of security of identification schemes shown by Goldreich[5].

Gabber *et al.*[4] 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, weak-unlinkability, and memory efficiency. However, the property of impersonation resistance was not much treated in this paper. Juang[8] and Hwang and Shiau[7] proposed authentication systems in multi-service environment with smart cards which satisfy memory efficiency. However, these systems cannot achieve weak-unlinkability. Liao and Wang[9] proposed the anonymous authentication system in multi-service environment with smart cards which have memory efficiency and weak-

unlinkability. However, service providers cannot identify each user in the system. Similarly, in anonymous credential systems[2] and in the systems based on group signatures[3], service providers cannot identify each user.

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, weak-unlinkability, and memory efficiency. In this section, we first show the definition of identification schemes[5]. Next, we discuss the extension of the definition of identification schemes based on an equality 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 $n \in \mathbf{N}$.

We show the definition of identification schemes. An identification scheme consists of a probabilistic algorithm and a protocol. Identification schemes are required that an entity (prover) is always able to convince another entity (verifier), whereas any adversary cannot fool the verifier into believing that he/she is the prover. Furthermore, any adversary cannot impersonate the prover even after receiving polynomially many proofs of identity from the prover.

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 a sequence of outputs $\langle \mathcal{P}(S), \mathcal{B}'(z) \rangle(\alpha, \mathcal{I}_S(\alpha))$ in polynomially many trials.

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 Protocols

In this section, we extend the identification scheme by an equality of protocols. We also prove that the extended identification schemes satisfies viability and impersonation resistance.

For any protocol $\langle \mathcal{A}, \mathcal{B} \rangle$ and any input x, 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, there exists a protocol $\langle \mathcal{A}'', \mathcal{B}'' \rangle$ such that
$$\langle \mathcal{A}'(x), \mathcal{B}'(x) \rangle = \langle \mathcal{A}''(x), \mathcal{B}'' \rangle.$$

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 $\langle \mathcal{P}', \mathcal{V}' \rangle$ in the previous lemma can be constructed from $\langle \mathcal{P}, \mathcal{V} \rangle$ 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 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 a sequence of outputs $\langle \mathcal{P}'(S), \mathcal{B}'(z) \rangle(\alpha, \mathcal{I}_S(\alpha))$ in polynomially many trials.

The proof can be found in the full-version of the paper.

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 a sequence of outputs $\langle \mathcal{P}'(S), \mathcal{B}'(z) \rangle(\alpha, \mathcal{I}_S(\alpha))$ in polynomially many trials.

3. Extension of Identification Scheme for Multi-Service Environment and Weak-unlinkability

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

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*. A user ID and a service ID 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 and G be multi-sets of functions mapping n-bit strings to n-bit strings and $|F| = |G| = 2^n$. The sets F and G are indexed by n-bit strings. We denote that $F = \{f_x : \{0,1\}^n \to \{0,1\}^n\}_{x \in \{0,1\}^n}$ and $G = \{g_y : \{0,1\}^n \to \{0,1\}^n\}_{y \in \{0,1\}^n}$. Here the indice x and y indicate user IDs. The preimages of functions belonging to F and G indicate a set of service IDs. The mapping of functions belonging to F indicates a set of secret-keys and that of functions belonging to G indicates a set of pseudonyms. That is, 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 $\langle \mathcal{P}, \mathcal{V} \rangle$ and multi-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 Weak-unlinkability

We define the property concerning privacy protection, which is called *weak-unlinkability*. 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 *weak-unlinkability* if for any probabilistic polynomial-time algorithm $\mathcal{A}$, any suffi-

ciently 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 $a \in \{0,1\}^n$ 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 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 weak-unlinkability.

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

In this section, we modify an identification scheme in multi-service environment which realizes an authentication system which satisfies impersonation resistance against insiders, weak-unlinkability, memory efficiency on auxiliary input tape, and personalization by using pseudorandom functions[6]. In the authentication system, a user behaves according to the one ITM in the protocol of the modified identification scheme and a service provider behaves according to the other ITM.

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.

4.1.1 Re-extended Identification Scheme

For any function f, let $\langle f \rangle$ be the description of an algorithm which on an input x returns $f(x)$. Then 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 $\mathcal{P}'$.

A *re-extended identification scheme* is a quadruplet of a verifying-key generating algorithm $\mathcal{I}$, a re-extended identification protocol $\langle \mathcal{P}'', \mathcal{V}' \rangle$, and sets of functions F and G. Our identification scheme is the re-extended identification scheme where F and G are pseudorandom functions. In what follows, we prove that our identification scheme satisfies impersonation resistance against insiders, weak-unlinkability, memory efficiency on auxiliary input tape, and personalization.

4.1.2 Pseudorandom Functions

A *pseudorandom function* is a multi-set of functions that map strings to strings and 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 weak-unlinkability 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 $\{0,1\}^n$.

The proof can be found in the full-version of the paper.

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 prove that our proposed scheme using pseudorandom functions as F and G satisfies impersonation resistance against insiders.

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

Theorem 2 If F and G are pseudorandom functions, for any identification scheme in multi-service environment $(\mathcal{I}, \langle \mathcal{P}, \mathcal{V} \rangle, F, G)$, 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 a sequence of outputs $\langle \mathcal{P}(f_U(b)), \mathcal{B}'(z) \rangle(g_U(b), \mathcal{I}_{f_U(b)}(g_U(b)))$ in polynomially many trials.

The proof can be found in the full-version of the paper. In the case where the common tape includes b in addition, it can be proved that the scheme has impersonation resistance against insiders.

The next theorem follows from Theorem 1 and Theorem 2.

Theorem 3 If F and G are pseudorandom functions, then any re-extended identification scheme $(\mathcal{I}, \langle \mathcal{P}'', \mathcal{V}' \rangle, F, G)$ 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 a sequence of outputs $\langle \mathcal{P}''(\langle f_a \rangle, \langle g_a \rangle), \mathcal{B}'(z) \rangle(b, \mathcal{I}_{f_U(b)}(g_U(b)))$ in polynomially many trials.

4.3 Evaluation of Weak-unlinkability

In this section, we prove that our proposed identification schemes satisfies weak-unlinkability.

Theorem 4 If F and G are pseudorandom functions, then any re-extended identification scheme $(\mathcal{I}, \langle \mathcal{P}'', \mathcal{V}' \rangle, F, G)$ has weak-unlinkability.

The proof can be found in the full-version of the paper.

4.4 Memory Efficiency on Auxiliary Input Tape

In our authentication system, a user behaves according to the one ITM $\mathcal{P}''$ in a re-extending identification scheme. Assuming that smart cards are used for authentication in our system, the auxiliary input tape of $\mathcal{P}''$ of our proposed identification scheme corresponds to the memory of each smart card in our authentication system. The memory efficiency on auxiliary input tape of identification schemes is defined as follows:

Definition 4 A re-extended identification scheme $(\mathcal{I}, \langle \mathcal{P}'', \mathcal{V}' \rangle, F, G)$ 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.

In a precise sense, the length of a service ID logarithmically depends on the number of service providers. However, the length of a service ID can be regarded as a constant since the size of a set of service IDs is generally much larger than the number of service providers. (In fact, all people in the world can even be indexed by only 40 bit strings.) Hence the algorithms which compute f_a and g_a do not depend on the number of service providers and the length of a service ID is independent of the number of service providers. Therefore, we have the following theorem.

Theorem 5 If the length of a service ID is a constant number, any re-extended identification scheme has memory efficiency on auxiliary input tape.

4.5 Personalization

The property of personalization is defined as follows:

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

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

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 If F and G are pseudorandom functions, then any re-extended identification scheme $(\mathcal{I}, \langle \mathcal{P}'', \mathcal{V}' \rangle, F, G)$ has personalization.

5. Conclusions

In this paper we proposed an authentication system in multi-service environment which satisfies impersonation resistance, weak-unlinkability, 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.

Our authentication system can be implemented based on the Schnorr identification scheme[13] and a collision-free hash function instead of pseudorandom functions. Our future work includes the implementation with smart cards and a PC in order to measure the execution time and to compare it with that of other related authentication systems.

Acknowledgements

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

References

1) Alsaid, A. and Martin, D.: Detecting Web Bugs with Bugnosis: Privacy Advocacy through Education, *Privacy Enhancing Technologies (PET2003)*, LNCS, Vol.2482, Springer-Verlag, pp.27–31 (2003).
2) Camenisch, J. and Lysyanskaya, A.: An Efficient System for Non-transferable Anonymous Credentials with Optional Anonymity Revocation, *Advances in Cryptology - EUROCRYPT 2001*, LNCS, Vol.2045, Springer-Verlag, pp.93–118 (2001).
3) Chaum, D. and van Heyst, E.: Group Signatures, *Advances in Cryptology - EUROCRYPT 1991*, LNCS, Vol.547, Springer-Verlag, pp.257–270 (1991).
4) Gabber, E., Information, C., Gibbons, P.B., Matias, Y. and Mayer, A.: How to Make Personalized Web Browsing Simple, Secure, and Anonymous, *Financial Cryptography*, LNCS, Vol.1318, Springer-Verlag, pp.17–31 (1997).
5) Goldreich, O.: *Foundations of Cryptography*, Cambridge University (2001).
6) Goldreich, O., Goldwasser, S. and Micali, S.: How to construct random functions, *Journal of the ACM (JACM)*, Vol.33, No.4, pp.792–807 (1986).
7) Hwang, R.-J. and Shiau, S.-H.: Provably Efficient Authenticated Key Agreement Protocol for Multi-Servers, *The Computer Journal*, Vol.50, pp.602–615 (2007).
8) Juang, W.-S.: Efficient multi-server password authenticated key agreement using smart cards, *IEEE Transactions on Consumer Electronics*, Vol.50, pp.251–255 (2004).
9) Liao, Y.-P. and Wang, S.-S.: A secure dynamic ID based remote user authentication scheme for multi-server environment, *Computer standards and interfaces*, Vol.31, No.1, pp.24–29 (2009).
10) OpenID: `http://openid.net/`.
11) OpenIDeas, the JanRain Blog, July 1, 2009: `http://blog.janrain.com/2009_07_01_archive.html`.
12) OpenID.ee: `https://openid.ee/en/`.
13) Schnorr, C.P.: Efficient signature generation by smart cards, *Journal of Cryptology*, Vol.4, No.3, Springer New York, pp.161–174 (1991).
14) VISA Corporate: `http://corporate.visa.com/_media/visa-corporate-overview.pdf`.